GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2002, 11:06:26 ; Search time 3842.15 Seconds
                                                 (without alignments)
                                                 1854.894 Million cell updates/sec

Title: US-09-394-745-6514
Perfect score: 432
Sequence: 1 gtccagcagctcggacttac..........attttctttttttttcttgg 432

Scoring table: IDENTITY_NUC
                   Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
                   1: gb_ba:*
                   2: gb_htg:*
                   3: gb_in:*
                   4: gb_om:*
                   5: gb_ov:*
                   6: gb_pat:*
                   7: gb_ph:*
                   8: gb_pl:*
                   9: gb_pr:*
                   10: gb_ro:*
                   11: gb_sts:*
                   12: gb_sy:*
                   13: gb_un:*
                   14: gb_vi:*
                   15: em_ba:*
                   16: em_fun:*
                   17: em_hum:*
                   18: em_in:*
                   19: em_om:*
                   20: em_or:*
                   21: em_ov:*
                   22: em_pat:*
                   23: em_ph:*
                   24: em_pl:*
                   25: em_ro:*
                   26: em_sts:*
                   27: em_sy:*

```
28:  em_un:*
29:  em_vi:*
30:  em_htgo_hum:*
31:  em_htgo_inv:*
32:  em_htgo_rod:*
33:  em_htg_hum:*
34:  em_htg_inv:*
35:  em_htg_rod:*
36:  em_htg_other:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	139.4	32.3	128017	8	AC084282	AC084282 Oryza sat
	2	138.4	32.0	1775	8	AF321856	AF321856 Lolium ri
	3	136.2	31.5	1810	8	AF321857	AF321857 Lolium ri
	4	133	30.8	1795	8	AF321855	AF321855 Lolium ri
	5	81.2	18.8	433	8	AF140486	AF140486 Oryza sat
	6	69	16.0	1373	8	AF082028	AF082028 Hemerocal
c	7	68.8	15.9	165909	2	AP003711	AP003711 Oryza sat
	8	60.6	14.0	1652	8	D78607	D78607 Arabidopsis
	9	59.6	13.8	622	6	AR074108	AR074108 Sequence
c	10	59.2	13.7	101009	8	ATF6G17	AL035601 Arabidops
c	11	59.2	13.7	196339	8	ATCHRIV87	AL161591 Arabidops
	12	59.2	13.7	206420	8	ATAP21	Z99707 Arabidopsis
	13	59	13.7	1656	8	AY039844	AY039844 Arabidops
	14	59	13.7	101009	8	ATF6G17	AL035601 Arabidops
	15	59	13.7	196339	8	ATCHRIV87	AL161591 Arabidops
	16	56.8	13.1	769	8	AF088221	AF088221 Oryza sat
	17	56.6	13.1	1674	8	HTCYP81C	AJ000477 Helianthu
	18	56.6	13.1	1719	8	HTCYP81L	AJ000478 Helianthu
c	19	54.6	12.6	138858	8	AP002968	AP002968 Oryza sat
c	20	54.6	12.6	156393	8	AP003204	AP003204 Oryza sat
	21	54.4	12.6	1806	8	AB001379	AB001379 Glycyrrhi
	22	53.2	12.3	4352	8	ZMCP71C1G	X81828 Z.mays CYP7
	23	52.6	12.2	72415	2	H0102C09	AL442103 Oryza sat
c	24	52.4	12.1	163055	2	AP003626	AP003626 Oryza sat
	25	51.6	11.9	1185	8	AF004210	AF004210 Zea mays
	26	51.6	11.9	1890	8	ZMCYP71C1	X81827 Z.mays CYP7
c	27	50.8	11.8	63773	8	AC074025	AC074025 Arabidops
	28	50.8	11.8	124253	2	AP003990	AP003990 Oryza sat
c	29	50.8	11.8	130451	2	AP004000	AP004000 Oryza sat
	30	50	11.6	161185	2	AP003571	AP003571 Oryza sat
	31	49.6	11.5	103463	8	AP003434	AP003434 Oryza sat
	32	49.4	11.4	1820	8	AB022732	AB022732 Glycyrrhi
	33	49.2	11.4	107491	8	AC002391	AC002391 Arabidops
	34	48	11.1	1711	8	CAR239051	AJ239051 Cicer ari
	35	47.8	11.1	1500	8	AB025016	AB025016 Lotus jap
c	36	47.4	11.0	100267	2	AP004022	AP004022 Oryza sat
c	37	47.4	11.0	116790	2	AP003976	AP003976 Oryza sat
	38	47.2	10.9	174301	2	AP003523	AP003523 Oryza sat

	39	47	10.9	171593	2	AP003522	AP003522 Oryza sat
	40	46.2	10.7	1793	8	ZMCYP71C3	X81830 Z.mays CYP7
	41	46.2	10.7	5057	8	ZMCI31AC3	Y11403 Z.mays cyp7
	42	45.4	10.5	152172	8	AC068924	AC068924 Oryza sat
c	43	44.4	10.3	73391	8	AP002057	AP002057 Arabidops
	44	43.8	10.1	1580	8	D78606	D78606 Arabidopsis
c	45	43.8	10.1	76994	8	AB026661	AB026661 Arabidops

ALIGNMENTS

RESULT   1
AC084282
LOCUS   AC084282   128017 bp   DNA   PLN   19-JUN-2001
DEFINITION Oryza sativa chromosome 3 BAC OSJNBb0048A17 genomic sequence,
            complete sequence.
ACCESSION   AC084282
VERSION   AC084282.6 GI:14389338
KEYWORDS   HTG.
SOURCE   Oryza sativa.
  ORGANISM Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 128017)
  AUTHORS   Buell,C.R., Yuan,Q., Ouyang,S., Moffat,K.S., Hill,J.N.,
            Gansberger,K., Brenner,M., Burgess,S., Hance,M., Shvartsbeyn,M.,
            Tsitrin,T., Riggs,F., Hsiao,J., Zismann,V., Blunt,S., Pai,G.,
            VanAken,S.E., Utterback,T.R., Feldblyum,T.V., Quackenbush,J.,
            Salzberg,S.L., White,O. and Fraser,C.M.
  TITLE   Oryza sativa chromosome 3 BAC OSJNBb0048A17 genomic sequence
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 128017)
  AUTHORS   Buell,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (20-OCT-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
REFERENCE   3 (bases 1 to 128017)
  AUTHORS   Buell,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (13-JUN-2001) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
REFERENCE   4 (bases 1 to 128017)
  AUTHORS   Buell,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-JUN-2001) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
COMMENT   On Jun 13, 2001 this sequence version replaced gi:12039441.
            Address all correspondence to:rice@tigr.org

            BAC clone OSJNBb0048A17 is from Oryza sativa chromosome 3
            The orientation of the sequence is from SP6 to T7 end of the BAC
            clone.
            Genes were identified by a combination of several methods: Gene
            prediction programs including Fgenesh (http://www.softberry.com/),
            genscan and Genscan+ (Chris Burge,

Matches 191; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

```
Qy     61 agcaagagctctggatggtcattagcatgtcctctgttgcggtcgtgaagttcttcctca 120
             |  |||     |  ||| ||   |   || ||||    |  ||    |  || ||||||| |   |  |||
Db      5 AACAACTAATATGGTTGTACTCTATCATGATATTTGCAACTGTGGTGAAGCTTATACTCT 64

Qy    121 tgctctactgccgaacgttcaagaatgagatcgtgagggcctacgcccaggaccatttct 180
             ||| ||||||| |||   |    ||    |||| | ||||| ||   | || ||     ||
Db     65 GGCTTTACTGCAGAAGCTCGAGAAACAAGATTGTCCGTGCCTATGCAGATGATCACCACT 124

Qy    181 tcgacgtaatcacaaactctgtcggcctggtctcggcgctgctcgctgtccggtacaaat 240
             | || ||   |  |||||      || ||   | ||   ||||   | || | ||      |   | |
Db    125 TTGATGTGGTAACAAATGTAGTTGGATTAGTTGCGGCTATTCTTGGTGATAAATTTTACT 184

Qy    241 ggtggatggaccctgttggcgccatactgatcgcgttgtacacgatcacgacgtgggcgc 300
             ||||||| || ||  | || || ||  ||  | ||||| ||| ||| | |    ||| | |
Db    185 GGTGGATTGATCCGATAGGAGCTATTTTGCTTGCAATTTACACCATCTCAAATTGGTCTC 244

Qy    301 gaacggtgctggagaacgtaggcacactgataggcaagtcggcgccggcagagtacctga 360
             | || ||  |||||||  |  |    || || ||  || || || | ||               ||
Db    245 GCACTGTCATGGAGAATGCCGTTTCATTGGTGGGACAATCTGCACCTCCTGAAGTTTTGC 304

Qy    361 cgaagctcacgtacttgatctggaaccacc 390
             |||||| || ||   |   |        |   ||||
Db    305 AGAAGCTAACATATCTCGTTATAAGGCACC 334
```


Search completed: February 7, 2002, 08:21:05
Job time: 18142 sec

http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHMM (Mark Borodovsky,
          http://genemark.biology.gatech.edu/GeneMark/), and GeneSplicer
          (Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org),
          searches of the complete sequence against a peptide database and
          the plant EST database at TIGR (http://www.tigr.org/tdb/tgi.shtml).
          Annotated genes are named to indicate the level of evidence for
          their annotation. Genes with similarity to other proteins are named
          after the database hits. Genes without significant peptide
          similarity but with EST similarity are named as unknown proteins.
          Genes without protein or EST similarity, that are predicted by more
          than two gene prediction programs over most of their length are
          annotated as hypothetical proteins. Genes encoding tRNAs are
          predicted by tRNAscan-SE (Sean Eddy,
          http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
          identified by repeatmasker (Arian Smit,
          http://ftp.genome.washington.edu/RM/RepeatMasker.html).
FEATURES   Location/Qualifiers
     source   1. .128017
                     /organism="Oryza sativa"
                     /cultivar="Nipponbare"
                     /sub_species="japonica"
                     /db_xref="taxon:4530"
                     /chromosome="3"
                     /map="R2404"
                     /clone="OSJNBb0048A17"
     repeat_region   complement(4411. .4455)
                     /rpt_family="AT_rich"
     mRNA   join(5116. .5166,5274. .5423,6406. .6528,6614. .6670,
                     6955. .7029,7121. .7523)
                     /gene="OSJNBb0048A17.2"
     gene   5116. .7523
                     /gene="OSJNBb0048A17.2"
                     /note="nearly identical to translation initiation factor
                     5A GB:CAB96075 GI:8919176 (Oryza sativa); EST AU057661,
                     AU108424 from this gene"
     CDS   join(5298. .5423,6406. .6528,6614. .6670,6955. .7029,
                     7121. .7225)
                     /gene="OSJNBb0048A17.2"
                     /codon_start=1
                     /product="translation initiation factor 5A"
                     /protein_id="AAK63944.1"
                     /db_xref="GI:14488377"
                     /translation="MSDSEEHHFESKADAGASKTYPQQAGTIRKNGHIVIKNRPCKVV
                     EVSTSKTGKHGHAKCHFVAIDIFNGKKLEDIVPSSHNCDVPHVNRTDYQLIDISEDGF
                     VSLLTESGGTKDDLRLPSDEALLTQIKDGFAEGKDLIVTVMSAMGEEQICALKDIGPK
                     N"
     repeat_region   complement(7835. .7913)
                     /rpt_family="AT_rich"
     repeat_region   complement(7915. .8003)
                     /rpt_family="Gaigin_Ol2 MITE element from gb:U72728 Oryza
                     longistaminata receptor-like kinase protein (Xa21), family
                     member F, pseudogene sequence (233 to 384) 152 nt"
     repeat_region   complement(7921. .8041)
                     /rpt_family="Gaijin_Os3 element from gb:D32165 Rice gene
                     for aspartic protease (302 to 448) 147 nt"
     mRNA   join(<8181. .8441,8581. .8675,8890. .9007,9093. .9211,
                     9300. .9486,9573. .9719,10135. .10681)

```
                          /gene="OSJNBb0048A17.3"
       gene               8181. .10681
                          /gene="OSJNBb0048A17.3"
                          /note="EST AU029806 from this gene"
       CDS                join(8181. .8441,8581. .8675,8890. .9007,9093. .9211,
                          9300. .9486,9573. .9719,10135. .10425)
                          /gene="OSJNBb0048A17.3"
                          /codon_start=1
                          /product="unknown protein"
                          /protein_id="AAK63921.1"
                          /db_xref="GI:14488354"
                          /translation="MAEALVAVLRLAASAAATARPQSRSGRHGSCAARVPCPGPSPFR
                          RGRLCARAAVAGPPEVDDDDAMTIDNLRRFFDVNVGKWNGAFYQFDAHGRVLQGISTR
                          LSVSTYGEDDLISLLQSLYIKQASSQISFVDEEDSEEWVEYKIKETNMFTVDKYQQVG
                          FFQEEKAFALRYQTAGMLETVLRAGVLGEDDTGEESPKNLKIPSRKPSIVCENCLYSR
                          EGNGRVRAFHIMDPKGVLDMLIIFHEKQGSEVPLMYSSDDADITNSDRIAPLLGRWEG
                          RSVTKRSGVYGATLSEADTVVLLEKDRNGQLILDNMSTKSGSSTTTTVHWTGSANNNL
                          LQFDGGYEMTLLPGGMYMGYPTDIGKIVNDMDSFHLEFCWMESPGKRQRLVRTYDSAG
                          LAVSSTYFFETKV"
       repeat_region      complement(8326. .8355)
                          /rpt_family="GC_rich"
       repeat_region      complement(8464. .8503)
                          /rpt_family="(GAA)n"
       repeat_region      12217. .12281
                          /rpt_family="(CGG)n"
       mRNA               join(<12233. .12508,12919. .13577,13715. .13858,
                          15212. .>15452)
                          /gene="OSJNBb0048A17.12"
       gene               12233. .15452
                          /gene="OSJNBb0048A17.12"
                          /note="similar to DNA binding protein GB:CAA88326
                          GI:1159877 (Avena fatua)"
       CDS                join(12233. .12508,12919. .13577,13715. .13858,15212.
 .15452)
                          /gene="OSJNBb0048A17.12"
                          /codon_start=1
                          /product="putative DNA-binding protein"
                          /protein_id="AAK63923.1"
                          /db_xref="GI:14488356"
                          /translation="MADRRRSDGGGGMQQQPFTSPGQERVFDGGGVPGQVAAPYGSDF
                          DQSSYMALLAAGAVGVGVGVQPTAAPWAVEEDVAAAPPGISLAPQFSMANYAPPPSYQ
                          HPATLVSPPLAAGLHPYPPYLHGVDAPPPQWPPRPAPPPSFSVLDLAAAAAPHEQRHS
                          MQQLLLRAAAFGGGMHAAAAPAPAAAAAIEQPAKDGYNWRKYGQKQLKDAESPRSYYK
                          CTRDGCPVKKIVERSSDGCIKEITYKGRHSHPRPVEPRRGGAASSSSSAMAAGTDHNA
                          GAAADDAAAADEDDPSDDDDTLLHEDDDDGEEGHDRGVDGEVGQRVVRKPKIILQTRS
                          EVDLLDDGYRWRKYGQKVVKGNPRPRSYYKCTADGCNVRKQIERASADPKCVLTTYTG
                          RHNHDPPGRPPAAANLQMPGPAAMRLAGGGTAHQQPSGGAHQMKEET"
       repeat_region      complement(12929. .13100)
                          /rpt_family="(CGG)n"
       repeat_region      13144. .13205
                          /rpt_family="(CGG)n"
       repeat_region      complement(13404. .13425)
                          /rpt_family="GC_rich"
       repeat_region      13463. .13565
                          /rpt_family="(CGA)n"
       repeat_region      complement(14525. .14573)
                          /rpt_family="AT_rich"
```

```
       repeat_region       complement(15595. .15615)
                           /rpt_family="(CAA)n"
       repeat_region       complement(16115. .16162)
                           /rpt_family="(GA)n"
       mRNA                complement(join(<16194.. .16307,16392. .16514,16642. .16836,
                           16923. .17048,17142. .17304,17388. .17523,17675. .18034,
                           18144. .18242,18340. .>18607))
                           /gene="OSJNBb0048A17.13"
       gene                complement(16194. .18607)
                           /gene="OSJNBb0048A17.13"
                           /note="predicted by fgenesh"
       CDS                 complement(join(16194. .16307,16392. .16514,16642. .16836,
                           16923. .17048,17142. .17304,17388. .17523,17675. .18034,
                           18144. .18242,18340. .18607))
                           /gene="OSJNBb0048A17.13"
                           /codon_start=1
                           /product="hypothetical protein"
                           /protein_id="AAK63926.1"
                           /db_xref="GI:14488359"
                           /translation="MGSCVSTTRRRRRSRKLSVAARKFRRKVSAAIADAPIARSGGGG
                           GAGGEVAAANCFARHEVVHVEAPVSNVTLHLTQLQWQHSQMDAGSVICEEAWYDSVSI
                           LDSADSEDDDLDNDFASVSGDPLPDVTATATSTSTSLLDAVHRLRSIASAEACQDDDP
                           PGKAEEESNAAAAADECCSSSGGGLKESAASSTRPPFPPSIPSNKIQPMPIVSVSPHSQ
                           KKKSAVVRLSFRRRSYEGDEMTEMSGSTNYLYRPRAGSSLPCSTGEKLSDGCWSAIEP
                           SVFRVRGESFFKDKRKSPAPNCSPYIPIGADMFACTRKINHIAQHLALPSLKAHETFP
                           SLLIVNIQMPTYPATVFGENDGDGISLVLYFKLSDSFDKEISPQLKESIKKLMGDEME
                           RVKGFPVDSNVPYTERLKILAGLVNPDDLQLSAAERKLVQTYNQKPVLSRPQHKFFKG
                           PNYFEIDLDVHRFSFISRKGLEAFRERLKHGVLDLGLTIQAQKAEELPEHVLCCMRLN
                           KIDFADSGQIPTLIMSSDE"
       repeat_region       complement(18449. .18488)
                           /rpt_family="(CGG)n"
       repeat_region       18751. .18773
                           /rpt_family="(CAAT)n"
       repeat_region       20529. .20560
                           /rpt_family="(GGAGAA)n"
       mRNA                join(<20577. .20761,21406. .21522,21610. .>22228)
                           /gene="OSJNBb0048A17.25"
       gene                20577. .22228
                           /gene="OSJNBb0048A17.25"
                           /note="predicted by genemarkHMM"
       CDS                 join(20577. .20761,21406. .21522,21610. .22228)
                           /gene="OSJNBb0048A17.25"
                           /codon_start=1
                           /product="hypothetical protein"

    Query Match            32.3%;   Score 139.4;   DB 8;   Length 128017;
    Best Local Similarity  71.4%;   Pred. No. 6.4e-25;
    Matches  182;  Conservative    0;  Mismatches  73;  Indels   0;  Gaps    0;

Qy      41 ggcggcaaggccaagggggcccctgctgatccctttcgggatggggcggcccaattgcccc 100
           ||||||  |   |  | ||    |||    ||| || |||||||||||||| |||   |  ||||||
Db   55653 GGCGGCGGGTGCGACGGCAACCTCTCGATGCCGTTCGGGATGGGGAGGCGGAGGTGCCCC 55712

Qy     101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160
           || || ||||||  ||  ||||    ||  ||  || |||||||||||||  |  ||||||| || | ||||||
Db   55713 GGCGAGACGCTGGCTCTGCACACGGTGGGGCTGGTGCTGGGCACGCTGATCCAGTGCTTC 55772
```

```
Qy    161  gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg 220
           |||||||| | ||| ||||| |    ||||       |  |    ||||||||| ||||||
Db  55773  GACTGGGAGAGGGTCGATGGCGTGGAGGTCGACATGGCTGAGGGTGGCGGGCTCACCATG 55832

Qy    221  ccccgggccgtcccgttggaggccatgtgcagccgcgtacagctatgcgtggtgttctt 280
           |||   || ||| |||||||||||||| |||||| ||||||    || ||| ||||||||||||
Db  55833  CCCAAGGTCGTGCCGTTGGAGGCCGTGTGCAGGCCGCGCGACGCCATGGGTGGTGTTCTT 55892

Qy    281  aagaggctctgaaaa 295
           ||||||||| |
Db  55893  CGCGAGCTCTGAACA 55907


RESULT    2
AF321856
LOCUS       AF321856    1775 bp    mRNA              PLN        18-APR-2001
DEFINITION  Lolium rigidum clone FHH-t putative cytochrome P450 mRNA, complete
            cds.
ACCESSION   AF321856
VERSION     AF321856.1  GI:13661745
KEYWORDS    .
SOURCE      Lolium rigidum.
  ORGANISM  Lolium rigidum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Poeae; Lolium.
REFERENCE   1  (bases 1 to 1775)
  AUTHORS   Fischer,T.C., Klattig,J.T. and Gierl,A.
  TITLE     A general cloning strategy of divergent plant cytochrome P450 genes
            and its application in Lolium rigidum and Ocimum basilicum
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1775)
  AUTHORS   Fischer,T.C., Klattig,J.T. and Gierl,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2000) Lehrstuhl fuer Zierpflanzenbau,
            TU-Muenchen, Am Hochanger 4, Freising 85350, Germany
FEATURES            Location/Qualifiers
     source         1. .1775
                    /organism="Lolium rigidum"
                    /isolate="SLR 31"
                    /db_xref="taxon:89674"
                    /clone="FHH-t"
     CDS            32. .1585
                    /codon_start=1
                    /product="putative cytochrome P450"
                    /protein_id="AAK38080.1"
                    /db_xref="GI:13661746"
                    /translation="MDKAYIAILSCAFLFLVHYVLGKVSDGRRGKKGAVQLPPSPPAV
                    PFLGHLHLVDKPIHATMCRLAARLGPVFSLRLGSRRAVVVSSSECARECFTEHDVTFA
                    NRPKFPSQLLVSFNGTALVTSSYGPHWRNLRRVATVQLLSAHRVACMSGVIAAEVRAM
                    ARRLFHATEASPDGAARVQLKRRLFELSLSVLMETIAQTKATRSEADADTDMSVEAQE
                    FKEVVDKLIPHLGAANMWDYLPVMRWFDVFGVRNKILHAVSRRDAFLRRLIDAERRRL
                    ADGGSDGDKKSMIAVLLTLQKTEPKVYTDTMITALCANLFGAGTETTSTTTEWAMSLL
                    LNHPAALKKAQAEIDASVGTSRLVSVDDVPSLAYLQCIVSETLRLYPAAPLLLPHESS
                    ADCKVGGYNVPADTMLIVNAYAIHRDPAAWEDPLEFRPERFEDGKAEGLFMIPFGMGR
                    RRCPGETLALRTIGMVLATLVQCFDWEPVDGVKVDMTEGGGFTIPKAVPLEAVCRPRA
                    VMRDVLQNL"
```

```
BASE COUNT       316 a      580 c      558 g      321 t
ORIGIN


   Query Match              32.0%;   Score 138.4;  DB 8;   Length 1775;
   Best Local Similarity    69.3%;   Pred. No. 1.5e-24;
   Matches  187;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;


Qy      41 ggcggcaaggccaaggggcccctgctgatccctttcgggatggggcggcccaattgcccc 100
           | |||||||||| |||| |    |    |||| || |||||||||||||||| |  ||||||
Db    1334 GACGGCAAGGCCGAGGGTCTGTTCATGATACCGTTCGGGATGGGGCGGCGGAGGTGCCCC 1393


Qy     101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160
           ||||| |||||| |||||||||||   |||| ||||  || || |||||| |  | ||||||
Db    1394 GGGGAGACGCTGGCGCTGCGGACGATCGGAATGGTCCTGGCGACGCTGGTGCAGTGCTTC 1453


Qy     161 gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg 220
           ||||||||| |||| || || |    ||||        |        || ||| | ||||||
Db    1454 GACTGGGAACCGGTGGACGGCGTGAAGGTGGACATGACGGAGGGGGGAGGGTTCACCATC 1513


Qy     221 ccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatgcgtggtgttctt 280
           ||     |||||| |||||||||||||| |||||| ||||||| | | |||||| | || |||
Db    1514 CCAAAGGCCGTGCCGTTGGAGGCCGTGTGCAGGCCGCGCGCGGTCATGCGCGACGTGCTT 1573


Qy     281 aagaggctctgaaaacctcatggatcgaat 310
           |||  |||| |  |  ||   |   | | ||
Db    1574 CAGAACCTCTAATCAACTAGTACCTTGCAT 1603



RESULT   3
AF321857
LOCUS       AF321857     1810 bp    mRNA              PLN       18-APR-2001
DEFINITION  Lolium rigidum clone FHH-y putative cytochrome P450 mRNA, complete
            cds.
ACCESSION   AF321857
VERSION     AF321857.1  GI:13661747
KEYWORDS    .
SOURCE      Lolium rigidum.
  ORGANISM  Lolium rigidum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Poeae; Lolium.
REFERENCE   1  (bases 1 to 1810)
  AUTHORS   Fischer,T.C., Klattig,J.T. and Gierl,A.
  TITLE     A general cloning strategy of divergent plant cytochrome P450 genes
            and its application in Lolium rigidum and Ocimum basilicum
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1810)
  AUTHORS   Fischer,T.C., Klattig,J.T. and Gierl,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2000) Lehrstuhl fuer Zierpflanzenbau,
            TU-Muenchen, Am Hochanger 4, Freising 85350, Germany
FEATURES            Location/Qualifiers
     source         1..1810
                    /organism="Lolium rigidum"
                    /isolate="SLR 31"
                    /db_xref="taxon:89674"
```

```
                         /clone="FHH-y"
            CDS          38. .1591
                         /codon_start=1
                         /product="putative cytochrome P450"
                         /protein_id="AAK38081.1"
                         /db_xref="GI:13661748"
                         /translation="MDKAYIAILSCAFLFLVHYVLGKVSDGRRGKKGAVQLPPSPPAV
                         PFLGHLHLVDKPIHATMCRLAARLGPVFSLRLGSRRAVVVSSSECARECFTEHDVTFA
                         NRPKFPSQLLVSFNGTALVTSSYGPHWRNLRRVATVQLLSAHRVACMSGVIAAEVRAM
                         ARRLFHAAEASPDGAARVQLKRRLFELSLSVLMETIAQTKATRSEADADTDMSVEAQE
                         FKEVVDKLIPHLGAANMWDYLPVMRWFDVFGVRNKILHAVSRRDAFLRRLIDAERRRL
                         ADGGSDGDKKSMIAVLLTLQKTEPKVYTDTMITALCANLFGAGTETTSTTTEWAMSLL
                         LNHPAALKKAQAEIDASVGTSRLVSVDDVPSLAYLQCIVNETLRLYPAAPLLLPHESS
                         ADCKVGGYNVPADTMLIVNAYAIHRDPAAWEHPLVFRPERFEDGKAEGLFMIPFGMGR
                         RRCPGETLALRTIGMVLATLVQCFDWEPVDGVNVDMTEGGGFTIPKAVPLEAVCRPRA
                         VMRDVLQSI"
BASE COUNT        329 a     578 c     567 g     336 t
ORIGIN


    Query Match              31.5%;   Score 136.2;   DB 8;   Length 1810;
    Best Local Similarity    69.6%;   Pred. No. 5.5e-24;
    Matches  183;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;


Qy      41 ggcggcaaggccaaggggcccctgctgatcccctttcgggatggggcggcccaattgcccc 100
           | ||||||||||| |||| |    |    |||| || |||||||||||||||||| | ||||||
Db    1340 GACGGCAAGGCCGAGGGTCTGTTCATGATACCGTTCGGGATGGGGCGGCGGAGGTGCCCC 1399


Qy     101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160
           || ||| |||||| ||||||||||| |||| |||| || || |||||| |  | ||||||
Db    1400 GGGGAGACGCTGGCGCTGCGGACGATCGGAATGGTCCTGGCGACGCTGGTGCAGTGCTTC 1459


Qy     161 gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg 220
           ||||||||  |||| || || ||   | ||        |       || ||| | |||||
Db    1460 GACTGGGAACCGGTGGACGGCGTGAATGTGGACATGACGGAGGGGGGAGGGTTCACCATC 1519


Qy     221 ccccgggccgtcccgttggaggccatgtgcagccgcgtacagctatgcgtggtgttctt 280
           ||    ||||||| |||||||||||||| |||||| ||||||  | |  ||||| |  || ||| |
Db    1520 CCAAAGGCCGTGCCGTTGGAGGCCGTGTGCAGGCCGCGCGCGGTCATGCGCGACGTGCTT 1579


Qy     281 aagaggctctgaaaacctcatgg 303
           ||||  ||| |  |  ||  |  |
Db    1580 CAGAGCATCTAATCAACTAGTAG 1602



RESULT    4
AF321855
LOCUS          AF321855      1795 bp     mRNA            PLN         18-APR-2001
DEFINITION     Lolium rigidum clone FHH-v putative cytochrome P450 mRNA, complete
               cds.
ACCESSION      AF321855
VERSION        AF321855.1   GI:13661743
KEYWORDS       .
SOURCE         Lolium rigidum.
  ORGANISM     Lolium rigidum
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
```

FEATURES   Location/Qualifiers
     source   1. .1795
                     /organism="Lolium rigidum"
                     /isolate="SLR 31"
                     /db_xref="taxon:89674"
                     /clone="FHH-v"
     CDS   32. .1585
                     /codon_start=1
                     /product="putative cytochrome P450"
                     /protein_id="AAK38079.1"
                     /db_xref="GI:13661744"
                     /translation="MDKAYIAILSSAFLFLVHYVLGKVSDGRRGKKGAVQLPPSPPAV
                     PFLGHLHLVEKPIHATMCRLAARLGPVFSLRLGSRRAVVVSSSECARECFTEHDVTFA
                     NRPKFPSQLLVSFNGTALVTSSYGPHWRNLRRVATVQLLSAHRVTCMSGVIAAEVRAM
                     ARRLFHAAEASPDGAARVQLKRRLFELSLSVLMETIAQTKATRSEADADTDMSLEAQE
                     FKEVVDKLIPHLGAANMWDYLPVMRWFDVFGVRSKILHAVSRRDAFLRRLINAERRRL
                     ADGGSDGDKKSMIAVLLTLQKTEPKVYTDTMITALCANLFGAGTETTSTTTEWAMSLL
                     LNHPAALKKAQAEIDASVGTSRLVSVDDVPSLAYLQCIVSETLRLYPAAPLLLPHESS
                     ADCKVGGYNVPADTMLIVNAYAIHRDPAAWEDPLEFKPERFEDGKAEGLFMIPFGMGR
                     RRCPGETLALRTIGMVLATLVQCFDWEPVDGVKVDMTEGGGFTIPKAVPLEAVCRPRV
                     VMRDVLQNL"
BASE COUNT   325 a   582 c   559 g   329 t
ORIGIN

     Query Match   30.8%;   Score 133; DB 8; Length 1795;
     Best Local Similarity   69.8%;   Pred. No. 3.5e-23;
     Matches 178; Conservative   0; Mismatches   77; Indels   0; Gaps   0;

Qy   44 ggcaaggccaaggggcccctgctgatccctttcgggatggggcggcccaattgccccggg 103
          ||||||||| |||||| | |||| || |||||||||||||||| | |||||||||
Db   1337 GGCAAGGCCGAGGGGCTGTTCATGATACCGTTCGGGATGGGGCGGCGGAGGTGCCCCGGA 1396

Qy   104 gaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttcgac 163
          || |||||| || || ||||| |||| |||| |||| || || |||||| | | |||||||||
Db   1397 GAGACGCTAGCACTACGGACGATCGGCATGGTCCTGGCGACGCTGGTGCAGTGCTTCGAC 1456

Qy   164 tgggacacggttgatggagctcaggtttgacatgaagctacggcgggctgaccatgccc 223
          ||||| |||| || || | |||| || || ||| | ||||| ||||| ||
Db   1457 TGGGAACCGGTGGACGGCGTGAAGGTGGACATGACAGAGGGGGGAGGGTTCACCATCCCA 1516

Qy   224 cgggccgtcccgttggaggccatgtgcagccgcgtacagctatgcgtggtgttcttaag 283
          |||||| |||||||||||| |||||| |||||| ||||||| | | ||||| | || || ||
Db   1517 AAGGCCGTGCCGTTGGAGGCCGTGTGCAGGCCGCGCGTGGTCATGCGCGACGTGCTTCAG 1576

Qy   284 aggctctgaaaacct 298

```
              |  ||||  |   |  ||
Db   1577 AACCTCTAATCATCT 1591


RESULT    5
AF140486
LOCUS        AF140486      433 bp    mRNA            PLN         11-MAY-1999
DEFINITION   Oryza sativa cytochrome P450 mRNA, partial cds.
ACCESSION    AF140486
VERSION      AF140486.1  GI:4768971
KEYWORDS     .
SOURCE       Oryza sativa.
  ORGANISM   Oryza sativa
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Ehrhartoideae; Oryzeae; Oryza.
REFERENCE    1  (bases 1 to 433)
  AUTHORS    Liu,J. and Yang,J.
  TITLE      Suppression subtractive hybridization (SSH) identified candidate
             genes that are differentially expressed at rice young panicle
  JOURNAL    Unpublished
REFERENCE    2  (bases 1 to 433)
  AUTHORS    Liu,J. and Yang,J.S.
  TITLE      Direct Submission
  JOURNAL    Submitted (05-APR-1999) Genetics, Institute of Genetics, No.220
             Handan Road, Shanghai 200433, China
FEATURES             Location/Qualifiers
     source          1. .433
                     /organism="Oryza sativa"
                     /db_xref="taxon:4530"
                     /tissue_type="panicle"
     CDS             <1. .320
                     /codon_start=3
                     /product="cytochrome P450"
                     /protein_id="AAD29699.1"
                     /db_xref="GI:4768972"
                     /translation="SMQRDPRVWEDPDKFIPERFKGFKVDRSGWMMPFGMGRRKCPGE
                     GLALRTVGMALGVMIQCFQWERLGKKKVDMSEGSGLTMPTAVPLMAMCLPRVEMESVL
                     KSL"
BASE COUNT      116 a      83 c     123 g   . 111 t
ORIGIN


   Query Match            18.8%;  Score 81.2;  DB 8;  Length 433;
   Best Local Similarity  59.1%;  Pred. No. 4e-10;
   Matches  137;  Conservative   0;  Mismatches   95;  Indels    0;  Gaps    0;

Qy     64 gctgatccctttcgggatggggcggcccaattgccccggggaaacgctcgcgctgcggac 123
           |  ||||  || ||||||  ||||||  |||  ||  ||||||||  |||    ||  || ||   ||||
Db     92 GATGATGCCCTTCGGTATGGGGAGGCGGAAGTGCCCCGGTGAAGGCCTTGCTCTTAGGAC 151

Qy    124 cgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacggttgatggagc 183
           ||  |||  |||  |||  |                ||  |   | |||||  | |||||    |  |  |
Db    152 GGTGGGGATGGCGCTAGGGGTTATGATACAATGCTTTCAGTGGGAGCGGCTCGGAAAGAA 211

Qy    184 tcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcccgttggaggc 243
           |||||            |              |||||  ||||||||   ||||||  ||  |     |||
```

```
Db    212 GAAGGTTGATATGAGTGAAGGTTCTGGGCTCACCATGCCTACGGCCGTGCCTCTCATGGC 271

Qy    244 catgtgcangccgcgtacagctatgcgtggtgttcttaagaggctctgaaaa 295
          |||||||   || |||    |  |||      ||| || || || |||| ||| 
Db    272 CATGTGCCTACCACGTGTGGAGATGGAGTCTGTGCTCAAAAGTCTCTAGAAA 323
```

RESULT   6
AF082028
LOCUS   AF082028   1373 bp   mRNA   PLN   15-JUL-1999
DEFINITION Hemerocallis hybrid cultivar senescence-associated protein 3 (SA3)
            mRNA, partial cds.
ACCESSION   AF082028
VERSION   AF082028.1 GI:3551949
KEYWORDS   .
SOURCE   Hemerocallis hybrid cultivar.
  ORGANISM Hemerocallis hybrid cultivar
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
            Hemerocallidaceae; Hemerocallis.
REFERENCE   1 (bases 1 to 1373)
  AUTHORS   Panavas,T., Pikula,A., Reid,P.D., Rubinstein,B. and Walker,E.L.
  TITLE   Identification of senescence-associated genes from daylily petals
  JOURNAL   Plant Mol. Biol. 40 (2), 237-248 (1999)
  MEDLINE   99339248
REFERENCE   2 (bases 1 to 1373)
  AUTHORS   Panavas,T., Pikula,A., Reid,P.D., Rubinstein,B. and Walker,E.L.
  TITLE   Direct Submission
  JOURNAL   Submitted (04-AUG-1998) Biology, University of Massachusetts,
            Morrill Science Center, Amherst, MA 01003, USA
FEATURES   Location/Qualifiers
     source   1. .1373
                     /organism="Hemerocallis hybrid cultivar"
                     /cultivar="Stella d'Oro"
                     /db_xref="taxon:80862"
                     /tissue_type="senescing petals"
     gene   <1. .1373
                     /gene="SA3"
     CDS   <1. .1121
                     /gene="SA3"
                     /function="putative cyt P450-containing fatty acid
                     hydroxylase"
                     /note="mRNA accumulates in senescing petals"
                     /codon_start=3
                     /product="senescence-associated protein 3"
                     /protein_id="AAC34853.1"
                     /db_xref="GI:3551950"
                     /translation="STEIFSPVRIRSLAAVRQEEVKLMITGILASTSTDNSVKVNMKV
                     VFSELMFNVIMKIIAGKRYFGVNTDSEVEEGQKFRVVFDEMFSTLEVASPQDFLPFLK
                     WFGFKRMENRLTKLAKELDQLFQKLIEERRSERGKVQSTVIDVLLSLQETDREQYSDK
                     LIKGMILSLIAAGTHTTAGTMEWAMSLLLNHPEALLKVRDEIDKKVGQDRLIDHSDLQ
                     NLSYLNNAIKESLRLFPTAPLLLAHESSAECTVGGFTIPSNTILFANAYALHRDPKVW
                     TDPVSFKPERFENNGQQGSRIYVPFGLGRRSCPGEGLATQVVGLALGTLIQCFEWDRN
                     GEEKVDMTDGSGLAMHMEKPLEAMCKPRQSIVDVINRL"
BASE COUNT   440 a   253 c   329 g   351 t
ORIGIN

```
Query Match              16.0%;  Score 69;  DB 8;  Length 1373;
Best Local Similarity    55.8%;  Pred. No. 4.3e-07;
Matches  129;  Conservative   0;  Mismatches  102;  Indels    0;  Gaps    0;


Qy     69 tccctttcgggatggggcggcccaattgccccggggaaacgctcgcgctgcggaccgtcg 128
          | || |||||| ||||| |||  |  || || || |||  ||| ||  ||     || |
Db    898 TGCCATTCGGGTTGGGGAGGCGGAGCTGTCCAGGTGAAGGGCTAGCAACGCAAGTTGTGG 957


Qy    129 ggctggtgctcgcaacgctgctcaattgcttcgactgggacacggttgatggagctcagg 188
          | |||    | |  || || |  | |||||||| |||||||     || || ||   |||
Db    958 GTTTGGCTTTGGGGACATTGATTCAATGCTTCGAGTGGGACCGAAATGGTGAAGAGAAGG 1017


Qy    189 tttgacatgaagctancggcgggctgaccatgccccgggccgtcccgttggaggccatgt 248
          |           |       |||||| ||||||     ||    || | ||||| ||||
Db   1018 TGGACATGACTGACGGATCAGGGCTCGCCATGCATATGGAAAAGCCTCTAGAGGCTATGT 1077


Qy    249 gcangccgcgtacagctatgcgtggtgttcttaagaggctctgaaaacctc 299
          ||| || ||  |  |||  || ||| | || |||||| |  ||| ||
Db   1078 GCAAACCTCGCCAAAGTATTGTTGATGTCATCAATAGGCTTTAGAAATTTC 1128



RESULT    7
AP003711/c
LOCUS        AP003711    165909 bp    DNA              HTG        31-MAY-2001
DEFINITION   Oryza sativa chromosome 6 clone P0417G12, *** SEQUENCING IN
             PROGRESS ***, in ordered pieces.
ACCESSION    AP003711
VERSION      AP003711.1  GI:14270111
KEYWORDS     HTG; HTGS_PHASE2.
SOURCE       Oryza sativa (cultivar:Nipponbare) DNA, clone:P0417G12.
  ORGANISM   Oryza sativa
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Ehrhartoideae; Oryzeae; Oryza.
REFERENCE    1  (sites)
  AUTHORS    Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE      Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
             clone:P0417G12
  JOURNAL    Published Only in DataBase (2001) In press
REFERENCE    2  (bases 1 to 165909)
  AUTHORS    Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE      Direct Submission
  JOURNAL    Submitted (30-MAY-2001) Takuji Sasaki, National Institute of
             Agrobiological Resources, Rice Genome Research Program; Kannondai
             2-1-2, Tsukuba, Ibaraki 305-8602, Japan
             (E-mail:tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
             Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT      NOTE: It currently consists of 1 contigs. Gaps between the contigs
             are represented as runs of N. The order of the pieces is believed
             to be correct as given, however the sizes of the gaps between them
             are based on estimates that have provided by the submitter. This
             sequence will be replaced by the finished sequence as soon as it is
             available and the accession number will be preserved.
             * NOTE: This is a 'working draft' sequence.
             * This sequence will be replaced
             * by the finished sequence as soon as it is available and
```

```
                    * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .165909
                     /organism="Oryza sativa"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:4530"
                     /chromosome="6"
                     /clone="P0417G12"
BASE COUNT     47417 a  34811 c  35629 g  47902 t    150 others
ORIGIN
```

```
  Query Match              15.9%;  Score 68.8;  DB 2;  Length 165909;
  Best Local Similarity    58.4%;  Pred. No. 3.4e-07;
  Matches 118;  Conservative    0;  Mismatches  84;  Indels    0;  Gaps     0;


Qy     61 cctgctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcgcgctgcg 120
          ||  ||| | ||  |||||||  ||||||||  |  |||||||||  |   |||| ||| |||
Db  84463 CCACCTGCTGCCGTTCGGGTCGGGGCGGCGGATCTGCCCCGGCGCGTCGCTGGCGATGCT 84404


Qy    121 gaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacggttgatgg 180
          |  ||   |  || ||| ||  |  || | | |||||||||| ||| | | || |  ||
Db  84403 GGTGGTGCAGGCGGCGCTGGCCGCCATGGTGCAGTGCTTCGAGTGGAGCCCCGTCGGCGG 84344


Qy    181 agctcaggttttgacatgaagctancggcgggctgaccatgccccgggccgtcccgttgga 240
          || | |||       || |          |||||||||  |||| |||       |||| | |
Db  84343 CGCGCCGGTGGACATGGAGGAGGGGCCCGGGCTGACGCTGCCGCGGAAGCGCCCGCTCGT 84284


Qy    241 ggccatgtgcangccgcgtaca 262
          ||    |   | |||||| | |
Db  84283 CTGCACCGTCTCGCCGCGGATA 84262
```

```
RESULT    8
D78607
LOCUS       D78607        1652 bp     mRNA              PLN       09-JUN-1998
DEFINITION  Arabidopsis thaliana mRNA for cytochrome P450 monooxygenase,
            complete cds, clone P450-66-8.
ACCESSION   D78607
VERSION     D78607.1  GI:3164143
KEYWORDS    .
SOURCE      Arabidopsis thaliana (strain:columbia) 7-d seedlings cDNA to mRNA,
            clone_lib:lambda ZAP II clone:P450-66-8.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1652)
  AUTHORS   Mizutani,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-DEC-1995) to the DDBJ/EMBL/GenBank databases.
            Masaharu Mizutani, International Research Laboratories, Ciba-Geigy
            (Japan), Bio-organics Department; 10-66 Miyuki-cho, Takarazuka,
            Hyogo 665, Japan (E-mail:masaharu.mizutani@jpta.mhs.ciba.com,
            Tel:0797-74-2464, Fax:0797-74-2455)
REFERENCE   2  (sites)
  AUTHORS   Mizutani,M., Ward,E. and Ohta,D.
```

FEATURES   Location/Qualifiers
     source   1. .1652
                     /organism="Arabidopsis thaliana"
                     /strain="columbia"
                     /db_xref="taxon:3702"
                     /clone="P450-66-8"
                     /clone_lib="lambda ZAP II"
                     /tissue_type="7-d seedlings"
     gene   51. .1553
                     /gene="CYP91A2"
     CDS   51. .1553
                     /gene="CYP91A2"
                     /codon_start=1
                     /product="cytochrome P450 monooxygenase"
                     /protein_id="BAA28539.1"
                     /db_xref="GI:3164144"
                     /translation="MLYFILLPLLFLVISYKFLYSKTQRFNLPPGPPSRPFVGHLHLM
                     KPPIHRLLQRYSNQYGPIFSLRFGSRRVVVITSPSLAQESFTGQNDIVLSSRPLQLTA
                     KYVVYNHTTVGTAPYGDHWRNLRRMCSQEILSSHRLIIFQHIRKDEILRMLTRLSRYT
                     QTSNESNDFTHIELEPLLSDLTFNNIVRMVTGKRYYGDDVNNKEEAELFKKLVYDIAM
                     YSGANHSADYLPILKLFGNKFEKEVKAIGKSMDDILQRLLDECRRDKEGNTMVNHLIS
                     LQQQQPEYYTDVIIKGLMMSMMLAGTETSAVTLEWAMANLLRNPEVLEKARSEIDEKI
                     GKDRLIDESDIAVLPYLQNVVSETFRLFPVAPFLIPRSPTDDMKIGGYDVPRDTIVMV
                     NAWAIHRDPEIWEEPEKFNPDRYNDGCGSDYYVYKLMPFGNGRRTCPGAGLGQRIVTL
                     ALGTLIQCFEWENVKGEEMDMSESTGLGMRKMDPLRAMCRPRPIMSKLLL"
     polyA_signal   1630. .1635
BASE COUNT   502 a   406 c   348 g   396 t
ORIGIN

     Query Match   14.0%;   Score 60.6; DB 8;   Length 1652;
     Best Local Similarity   53.2%; Pred. No. 5.5e-05;
     Matches 126; Conservative   0; Mismatches 111; Indels   0; Gaps   0;

Qy   32 gatggctccggcggcaaggccaaggggcccctgctgatccctttcgggatggggcggccc 91
           || || | ||| || | | | | | |||||| || || |||| || |||
Db   1299 GACGGATGCGGAAGCGATTACTATGTTTACAAGCTGATGCCGTTTGGGAATGGCCGGAGA 1358

Qy   92 aattgccccgggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctc 151
           | ||| |||||| | | | | | ||| || ||| ||| | |||| || |
Db   1359 ACTTGTCCCGGCGCCGGATTAGGTCAGAGGATTGTGACATTGGCGCTTGGAACGTTGATT 1418

Qy   152 aattgcttcgactgggacacggttgatggagctcaggtttgacatgaagctancggcggg 211
           | |||||| || |||||| | || | || | || || | | | |||
Db   1419 CAATGCTTTGAATGGGAGAATGTGAAAGGGGAAGAGATGGATATGTCTGAGAGTACTGGG 1478

Qy   212 ctgaccatgccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatg 268
           || |||| | | || || || |||||||||| | ||| | | ||||
Db   1479 TTGGGTATGCGTAAGATGGATCCTTTACGGGCCATGTGTAGGCCTAGGCCCATTATG 1535

RESULT   9

AR074108

    Query Match   13.8%;   Score 59.6;   DB 6;   Length 622;
    Best Local Similarity   54.1%;   Pred. No. 0.00011;
    Matches 119; Conservative   0; Mismatches 101; Indels   0; Gaps   0;

Qy   61 cctgctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcgcgctgcg 120
            || || | || || |||.||!||| || | ||| || || | || || |
Db   180 CCGACTATTGCCGTTTGGGATGGGGAGGAGAAGTTGTCCTGGTGCTGGCCTTGCCAATAG 239

Qy   121 gaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacggttgatgg 180
            || | |||| || || || |||| | | |||||||||| ||||| |||| |
Db   240 AGTGGTGAGCTTGGTCCTGGCGGCGCTTATTCAGTGCTTCGAATGGGAACGAGTTGGCGA 299

Qy   181 agctcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcccgttgga 240
            || ||| | || || || ||||| | | || |||||
Db   300 AGAATTGGTGGACTTGTCCGAGGGGACGGGACTCACAATGCCAAAGAGAGAGCCATTGGA 359

Qy   241 ggccatgtgcangccgcgtacagctatgcgtggtgttctt 280
            |||| |||||| ||||| | ||| | | | | ||
Db   360 GGCCTTGTGCAAAGCGCGTGAATGCATGATAGCTAATGTT 399

RESULT 10
ATF6G17/c

FEATURES   Location/Qualifiers
     source   1. .101009
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     misc_feature   1. .18845
                     /note="position 1-18845 overlap to EMBL accession Z99707;
                     please refer to this entry for analysis and annotation"
     gene   complement(join(18286. .18900,19005. .19391,19523. .20020))
                     /gene="F6G17.10"
     gene   18286. .20020
                     /gene="F6G17.10"
     CDS   complement(join(18286. .18900,19005. .19391,19523. .20020))
                     /gene="F6G17.10"
                     /note="similarity to cytochrome P450 monooxygenase,
                     Arabidopsis thaliana, D78606
                     Contains Cytochrome P450 cysteine heme-iron ligand
                     signature[FGLGRRACPG]"
                     /codon_start=1
                     /product="cytochrome p450-like protein"
                     /protein_id="CAB38203.1"
                     /db_xref="GI:4468802"
                     /translation="MEALMLIFTFCFIVLSLIFLIGRIKRKLNLPPSPAWALPVIGHL
                     RLLKPPLHRVFLSVSQSLGDAPIISLRLGNRLLFVVSSHSIAEECFTKNDVILANRQT
                     TISTKHISYGNSTVVSASYSEHWRNLRRIGALEIFSAHRLNSFSSIRRDEIRRLIGRL
                     LRNSSYGFTKVEMKSMFSDLTFNNIIRMLAGKCYYGDGKEDDPEAKRVRTLIAEAMSS
                     SGPGNAADYIPILTWITYSETRIKKLAGRLDEFLQGLVDEKREGKEKKENTMVDHLLC
                     LQETQPEYYMDRIIKGTMLSLIAGGTDTTAVTLEWALSSLLNNPEVLNKARDEIDRMI
                     GVDRLLEESDIPNLPYLQNIVSETLRLYPAAPMLLPHVASKDCKVGGYDMPRGTMLLT
                     NAWAIHRDPLLWDDPTSFKPERFEKEGEAKKLMPFGLGRRACPGSGLAQRLVTLSLGS
                     LIQCFEWERIGEEEVDMTEGPGLTMPKARPLEAMCRARDFVGKILPDSS"
     exon   complement(18286. .18900)
                     /gene="F6G17.10"
                     /number=1
     intron   complement(18901. .19004)
                     /number=1
     exon   complement(19005. .19391)
                     /gene="F6G17.10"
                     /number=2
     intron   complement(19392. .19522)
                     /number=2
     exon   complement(19523. .20020)
                     /gene="F6G17.10"

```
                          /number=3
gene                      21133. .22840
                          /gene="F6G17.20"
exon                      complement(21133. .21744)
                          /gene="F6G17.20"
                          /number=1
gene                      complement(join(21133. .21744,21842. .22225,22343. .22840))
                          /gene="F6G17.20"
CDS                       complement(join(21133. .21744,21842. .22225,22343. .22840))
                          /gene="F6G17.20"
                          /note="similarity to cytochrome P450, Glycyrrhiza
                          echinata, AB001379
                          Contains Cytochrome P450 cysteine heme-iron ligand
                          signature [FGLGRRACPG]
                          contains EST gb:AA586064, H76015, T41596, N38867"
                          /codon_start=1
                          /product="cytochrome P450-like protein"
                          /protein_id="CAB38204.1"
                          /db_xref="GI:4468803"
                          /translation="METKTLIFSILFVVLSLIYLIGKLKRKPNLPPSPAWSLPVIGHL
                          RLLKPPIHRTFLSLSQSLNNAPIFSLRLGNRLVFVNSSHSIAEECFTKNDVVLANRPN
                          FILAKHVAYDYTTMIAASYGDHWRNLRRIGSVEIFSNHRLNSFLSIRKDEIRRLVFRL
                          SRNFSQEFVKVDMKSMLSDLTFNNILRMVAGKRYYGDGVEDDPEAKRVRQLIADVVAC
                          AGAGNAVDYLPVLRLVSDYETRVKKLAGRLDEFLQGLVDEKREAKEKGNTMIDHLLTL
                          QESQPDYFTDRIIKGNMLALILAGTDTSAVTLEWALSNVLNHPDVLNKARDEIDRKIG
                          LDRLMDESDISNLPYLQNIVSETLRLYPAAPMLLPHVASEDCKVAGYDMPRGTILLTN
                          VWAIHRDPQLWDDPMSFKPERFEKEGEAQKLMPFGLGRRACPGSGLAHRLINLTLGSL
                          IQCLEWEKIGEEVDMSEGKGVTMPKAKPLEAMCRARPSVVKIFNESV"
intron                    complement(21745. .21841)
                          /number=1
exon                      complement(21842. .22225)
                          /gene="F6G17.20"
                          /number=2
intron                    complement(22226. .22342)
                          /number=2
exon                      complement(22343. .22840)
                          /gene="F6G17.20"
                          /number=3
gene                      23202. .25100
                          /gene="F6G17.30"
CDS                       complement(23202. .25100)
                          /gene="F6G17.30"
                          /note="similarity to various predicted proteins,
                          Arabidopsis thaliana
                          Contains Cytochrome c family heme-binding site signature
                          [CSDCHT]"
                          /codon_start=1
                          /product="putative protein"
                          /protein_id="CAB38205.1"
                          /db_xref="GI:4468804"
                          /translation="MASSPLLATSLPQNQLSTTATARFRLPPPEKLAVLIDKSQSVDE
                          VLQIHAAILRHNLLLHPRYPVLNLKLHRAYASHGKIRHSLALFHQTIDPDLFLFTAAI
                          NTASINGLKDQAFLLYVQLLSSEINPNEFTFSSLLKSCSTKSGKLIHTHVLKFGLGID
                          PYVATGLVDVYAKGGDVVSAQKVFDRMPERSLVSSTAMITCYAKQGNVEAARALFDSM
                          CERDIVSWNVMIDGYAQHGFPNDALMLFQKLLAEGKPKPDEITVVAALSACSQIGALE
                          TGRWIHVFVKSSRIRLNVKVCTGLIDMYSKCGSLEEAVLVFNDTPRKDIVAWNAMIAG
                          YAMHGYSQDALRLFNEMQGITGLQPTDITFIGTLQACAHAGLVNEGIRIFESMGQEYG
```

```
                     IKPKIEHYGCLVSLLGRAGQLKRAYETIKNMNMDADSVLWSSVLGSCKLHGDFVLGKE
                     IAEYLIGLNIKNSGIYVLLSNIYASVGDYEGVAKVRNLMKEKGIVKEPGISTIEIENK
                     VHEFRAGDREHSKSKEIYTMLRKISERIKSHGYVPNTNTVLQDLEETEKEQSLQVHSE
                     RLAIAYGLISTKPGSPLKIFKNLRVCSDCHTVTKLISKITGRKIVMRDRNRFHHFTDG
                     SCSCGDFW"
     exon            complement(23202. .25100)
                     /gene="F6G17.30"
                     /number=1
     gene            complement(23202. .25100)
                     /gene="F6G17.30"
     gene            30884. .32930
                     /gene="F6G17.40"
     exon            30884. .31206
                     /gene="F6G17.40"
                     /number=1
     CDS             join(30884. .31206,31359. .31460,31544. .32930)
                     /gene="F6G17.40"
                     /note="similarity to hypothetical protein, Glycine max,
                     PIR2:S17433
                     contains EST gb:T44614, AAA712644, AA395496"
                     /codon_start=1
                     /product="auxin-responsive GH3-like protein"
                     /protein_id="CAB38206.1"
                     /db_xref="GI:4468805"
                     /translation="MAVDSPLQSRMVSATTSEKDVKALKFIEEMTRNPDSVQEKVLGE
                     ILTRNSNTEYLKRFDLDGVVDRKTFKSKVPVVTYEDLKPEIQRISNGDCSPILSSHPI
                     TEFLTSSGTSAGERKLMPTIEEDLDRRQLLYSLLMPVMNLYVPGLDKGKGLYFLFVKS
                     ESKTSGGLPARPVLTSYYKSDHFKRRPYDPYNVYTSPNEAILCSDSSQSMYAQMLCGL
                     LMRHEVLRLGAVFASGLLRAISFLQNNWKELARDISTGTLSSRIFDPAIKNRMSKILT
                     KPDQELAEFLVGVCSQENWEGIITKIWPNTKYLDVIVTGAMAQYIPTLEYYSGGLPMA
                     CTMYASSESYFGINLKPMCKPSEVSYTIMPNMAYFEFLPHNHDGDGAAEASLDETSLV
                     ELANVEVGKEYELVITTYAGLYRYRVGDILRVTGFHNSAPQFKFIRRKNVLLSVESDK
                     TDEAELQKAVENASRLFAEQGTRVIEYTSYAETKTIPGHYVIYWELLGRDQSNALMSE
                     EVMAKCCLEMEESLNSVYRQSRVADKSIGPLEIRVVRNGTFEELMDYAISRGASINQY
                     KVPRCVSFTPIMELLDSRVVSAHFSPSLPHWSPERRR"
     intron          31207. .31358
                     /gene="F6G17.40"
                     /number=1
     exon            31359. .31460
                     /gene="F6G17.40"
                     /number=2
     intron          31461. .31543
                     /gene="F6G17.40"
                     /number=2
     exon            31544. .32930
                     /gene="F6G17.40"
                     /number=3
     gene            35807. .37359
                     /gene="F6G17.50"
     CDS             join(35807. .36127,36230. .36289,36724. .37359)
                     /gene="F6G17.50"
                     /note="strong similarity to cytochrome P450 monooxygenase
                     CYP91A2, Arabidopsis thaliana, D78607
                     Contains Cytochrome P450 cysteine heme-iron ligand
                     signature [FGNGRRSCPG]
                     contains EST gb:AA712784"
                     /codon_start=1
```

```
                         /product="cytochrome P450 monooxygenase-like protein"
                         /protein_id="CAB38207.1"
                         /db_xref="GI:4468806"
                         /translation="MVTGKRYYGDEVHNEEEANVFKKLVADINDCSGARHPGDYLPFM
                         KMFGGSFEKKVKALAEAMDEILQRLLEECKRDKDGNTMVNHLLSLQQNEPEYYTDVTI
                         KGLMLIFCFFGQLQILWFTNIETGWGMMIAGTDTSAVTLEWAMSSLLNHPEALEKAKL
                         EIDEKIGQERLIDEPDIANLPYLQNIVSETFRLYPAAPLLVPRSPTEDIKVGGYDVPR
                         GTMVMVNAWAIHRDPELWNEPEKFKPERFNGGEGGGRGEDVHKLMPFGNGRRSCPGAG
                         LGQKIVTLALGSLIQCFDWQKVNGEAIDMTETPGMAMRKKIPLSALCQSRPIMSKLQA
                         HLKG"
      exon               35807..36127
                         /gene="F6G17.50"
                         /number=1
      intron             36128..36229
                         /gene="F6G17.50"

   Query Match               13.7%;   Score 59.2;  DB 8;  Length 101009;
   Best Local Similarity     54.1%;   Pred. No. 9.2e-05;
   Matches  118;  Conservative   0;  Mismatches  100;  Indels   0;  Gaps   0;


Qy     54 aggggcccctgctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcg 113
             ||| |    ||| || || || ||||||||| || |    ||| || || |    ||| |
Db .11133 AGGCTCGAAAGCTAATGCCGTTTGGGATGGGACGACGAGCTTGTCCTGGAGCTGAGCTTG 11074


Qy    114 cgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacgg 173
            |  ||||    ||  | || |   || |        || |  | |  ||||| ||||| | |  |
Db  11073 GGAAGCGGTTAGTGAGCCTTGCTCTTGGGTGCTTGATTCAGTCTTTCGAGTGGGAGAGAG 11014


Qy    174 ttgatggagctcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcc 233
            ||| || ||  |  ||        |    || ||| | || ||||| || |   |
Db  11013 TTGGTGCAGAACTTGTGGACATGACTGAAGGCGAAGGGATCACTATGCCTAAAGCTACTC 10954


Qy    234 cgttggaggccatgtgcangccgcgtacagctatgcgt 271
            |||||    ||  |||||||| | | ||| |     | |   ||
Db  10953 CGTTGCGAGCTATGTGCAAGGCACGTGCCATTGTTGGT 10916



RESULT  11
ATCHRIV87/c
LOCUS       ATCHRIV87  196339 bp     DNA              PLN        16-MAR-2000
DEFINITION  Arabidopsis thaliana DNA chromosome 4, contig fragment No. 87.
ACCESSION   AL161591
VERSION     AL161591.2  GI:7270703
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 42610 to 143618; 123423 to 196339)
  AUTHORS   Rose,M., Hempel,S., Entian,K.-D., Mewes,H.W., Lemcke,K. and
            Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 196339)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
```

```
                    Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
                    lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
                    Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
                    Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
                    E-mail: michael.bevan@bbsrc.ac.uk
     COMMENT        Information on performance of analysis and a more detailed
                    annotation of this entry and other sequences of chromosomes 3, 4
                    and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
                    this fragment has an overlap with ATCHRIV86 at the 5' end and an
                    overlap with ATCHRIV88 at the 3' end.
     FEATURES             Location/Qualifiers
          source         1. .196339
                         /organism="Arabidopsis thaliana"
                         /variety="Columbia"
                         /db_xref="taxon:3702"
                         /chromosome="4"
          gene           6146. .7792
                         /gene="AT4g37210"
          exon           6146. .6474
                         /gene="AT4g37210"
                         /number=1
          CDS            join(6146. .6474,6702. .6919,7013. .7225,7313. .7403,
                         7489. .7792)
                         /gene="AT4g37210"
                         /note="intron number 1 is a special U12 intron
                         similarity to nuclear histone-binding protein N1/N2,
                         Xenopus laevis, PIR2:A25680"
                         /codon_start=1
                         /product="putative protein"
                         /protein_id="CAB80387.1"
                         /db_xref="GI:7270704"
                         /translation="MVEESASASEASVIQTLTEPATEIAQTLEPNLASIEATVESVVQ
                         GGTESTCNNDANNNNAADSAATEVCDEEREKTLEFAEELTEKGSVFLKENDFAEAVDC
                         FSRALEIRVAHYGELDAECINAYYRYGLALLAKAQAEADPLGNMPKKEGEVQQESSNG
                         ESLAPSVVSGDPERQGSSSGQEGSGGKDQGEDGEDCQDDDLSDADGDADEDESDLDMA
                         WKMLDIARVITDKQSTETMEKVDILCSLAEVSLEREDIESSLSDYKNALSILERLVEP
                         DSRRTAELNFRICICLETGCQPKEAIPYCQKALLICKARMERLSNEIKGASGSATSST
                         VSEIDEGIQQSSNVPYIDKSASDKEVEIGDLAGLAEDLEKKASKLNLSVH"
          intron         6475. .6701
                         /gene="AT4g37210"
                         /number=1
          exon           6702. .6919
                         /gene="AT4g37210"
                         /number=2
          intron         6920. .7012
                         /gene="AT4g37210"
                         /number=2
          exon           7013. .7225
                         /gene="AT4g37210"
                         /number=3
          intron         7226. .7312
                         /gene="AT4g37210"
                         /number=3
          exon           7313. .7403
                         /gene="AT4g37210"
                         /number=4
          intron         7404. .7488
```

```
                              /gene="AT4g37210"
                              /number=4
          exon                7489. .7792
                              /gene="AT4g37210"
                              /number=5
          gene                8876. .9739
                              /gene="AT4g37220"
          CDS                 join(8876. .9051,9141. .9248,9330. .9400,9486. .9739)
                              /gene="AT4g37220"
                              /note="strong similarity to cold acclimation protein
                              WCOR413, Triticum aestivum, PATCHX:G1657855
                              Contains Prokaryotic membrane lipoprotein lipid attachment
                              site AA147-157
                              contains EST gb:AW033651.1, W43270, AA650647, AI996990.1,
                              AA728669, T42795, Z37671, AI100742, T42949, AA040998,
                              AA395771, AA657303, T41871, T45633"
                              /codon_start=1
                              /product="cold acclimation protein homolog"
                              /protein_id="CAB80388.1"
                              /db_xref="GI:7270705"
                              /translation="MGRGEFLAMKTEENAANLINSDMNEFVAAAKKLVKDVGMLGGVG
                              FGTSVLQWAASIFAIYLLILDRTNWKTKMLTTLLVPYIFFTLPSVIFQFFSGDFGKWI
                              ALIAIIVRLFFPKEFPEWLEIPVALILIVVVSPSLIAWTLRESWVGAVICLVIACYLF
                              HEHIKASGGFKNSFTQKNGISNTIGIVALLVYPVWTIFFHIF"
          exon                8876. .9051
                              /gene="AT4g37220"
                              /number=1
          intron              9052. .9140
                              /gene="AT4g37220"
                              /number=1
          exon                9141. .9248
                              /gene="AT4g37220"
                              /number=2
          intron              9249. .9329
                              /gene="AT4g37220"
                              /number=2
          exon                9330. .9400
                              /gene="AT4g37220"
                              /number=3
          intron              9401. .9485
                              /gene="AT4g37220"
                              /number=3
          exon                9486. .9739
                              /gene="AT4g37220"
                              /number=4
          gene                10194. .11357
                              /gene="AT4g37230"
          exon                10194. .10398
                              /gene="AT4g37230"
                              /number=1
          CDS                 join(10194. .10398,11131. .11357)
                              /gene="AT4g37230"
                              /note="possible frameshift in DNA sequence at pos.
                              51188-51200, 5' part of gene couldn't be reconstructed,
                              possible pseudogene, no ATG
                              strong similarity to photosystem II oxygen-evolving
                              complex protein 1, spinach, PIR2:A23626
```

```
                          contains EST gb:Z34685"
                          /codon_start=1
                          /product="photosystem II oxygen-evolving complex like
                          protein (partial)"
                          /protein_id="CAB80389.1"
                          /db_xref="GI:7270706"
                          /translation="MTSLTYTLDEIEGPFEVDYAAVTVHNFLVGSVYRSCSRSSSSWH
                          RRGYFGPKSIPSAFTQGHVGNKSDQYQGYDNAVALPARGNNEELAKENNKITLSVTKS
                          NPESGEVIGAFESIQPSDTDLGATTPKDVKIQGIWYCQLDE"
intron                    10399. .11130
                          /gene="AT4g37230"
                          /number=1
exon                      11131. .11357
                          /gene="AT4g37230"
                          /number=2
exon                      18388. .18822
                          /gene="AT4g37240"
                          /number=1
gene                      18388. .18822
                          /gene="AT4g37240"
CDS                       18388. .18822
                          /gene="AT4g37240"
                          /note="contains EST gb:Z18456"
                          /codon_start=1
                          /product="putative protein"
                          /protein_id="CAB80390.1"
                          /db_xref="GI:7270707"
                          /translation="MEFANPVKVGYVLLKYPMCFICNSDDMDFDDAVAAISADEELQL
                          GQIYFALPLCWLRQPLKAEEMAALAVKASSALMRGGGGGCRRKCVEPIVSDKLRMRVG
                          SGDDTVGSGSGRRKVRNGDGGGSVSSSRRRKCYAAELSTIDE"
gene                      21559. .23955
                          /gene="AT4g37250"
exon                      complement(21559. .22253)
                          /gene="AT4g37250"
                          /number=1
gene                      complement(join(21559. .22253,22350. .23955))
                          /gene="AT4g37250"
CDS                       complement(join(21559. .22253,22350. .23955))
                          /gene="AT4g37250"
                          /note="similarity to protein kinase TMK1, receptor type
                          precursor - Arabidopsis thaliana, PIR1:JQ1674
                          contains EST gb:H76836"
                          /codon_start=1
                          /product="receptor kinase-like protein"
                          /protein_id="CAB80391.1"
                          /db_xref="GI:7270708"
                          /translation="MELISVIFFFFCSVLSSSALNSDGLVLMKFKSSVLVDPLSLLQT
                          WNYKHESPCSWRGISCNNDSKVLTLSLPNSQLLGSIPSDLGSLLTLQSLDLSNNSFNG
                          PLPVSFFNARELRFLDLSSNMISGEIPSAIGDLHNLLTLNLSDNALAGKLPTNLASLR
                          NLTVVSLENNYFSGEIPGGWRVVEFLDLSSNLINGSLPPDFGGYSLQYLNVSFNQISG
                          EIPPEIGVNFPRNVTVDLSFNNLTGPIPDSPVFLNQESNFFSGNPGLCGEPTRNPCLI
                          PSSPSIVSEADVPTSTPAIAAIPNTIGSNPVTDPNSQQTDPNPRTGLRPGVIIGIVVG
                          DIAGIGILAVIFLYIYRCKKNKIVDNNNNDKQRTETDTITLSTFSSSSSSPEESRRFR
                          KWSCLRKDPETTPSEEEDEDDEDEESGYNANQRSGDNKLVTVDGEKEMEIETLLKASA
                          YILGATGSSIMYKAVLEDGRVFAVRRLGENGLSQRRFKDFEPHIRAIGKLVHPNLVRL
                          CGFYWGTDEKLVIYDFVPNGSLVNPRYRKGGGSSSPYHLPWETRLKIAKGIARGLAYL
                          HEKKHVHGNLKPSNILLGHDMEPKIGDFGLERLLTGETSYIRAGGSSRIFSSKRYTTS
```

```
                    SREFSSIGPTPSPSPSSVGAMSPYCAPESFRSLKPSPKWDVYGFGVILLELLTGKIVS
                    VEEIVLGNGLTVEDGHRAVRMADVAIRGELDGKQEFLLDCFKLGYSCASPVPQKRPTM
                    KESLAVLERFHPNSSVIKSSSFHYGH"
       intron       complement(22254. .22349)
                    /number=1
       exon         complement(22350. .23955)
                    /gene="AT4g37250"
                    /number=2
       gene         34372. .35334
                    /gene="AT4g37260"
       CDS          34372. .35334
```

```
  Query Match              13.7%;   Score 59.2;  DB 8;   Length 196339;
  Best Local Similarity   54.1%;   Pred. No. 8.7e-05;
  Matches  118;  Conservative   0;  Mismatches  100;  Indels   0;  Gaps    0;
```

```
Qy     54 aggggcccctgctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcg 113
          |||   |    ||| || || || |||||||| || |     ||| || || |    ||| |
Db  53741 AGGCTCGAAAGCTAATGCCGTTTGGGATGGGACGACGAGCTTGTCCTGGAGCTGAGCTTG 53682


Qy    114 cgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacgg 173
          |  ||||   ||  | ||| |  ||| |          || |  | | ||||| |||||| |  |
Db  53681 GGAAGCGGTTAGTGAGCCTTGCTCTTGGGTGCTTGATTCAGTCTTTCGAGTGGGAGAGAG 53622


Qy    174 ttgatggagctcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcc 233
          ||| || || | || |           |    ||  ||| | || |||||     ||     |
Db  53621 TTGGTGCAGAACTTGTGGACATGACTGAAGGCGAAGGGATCACTATGCCTAAAGCTACTC 53562


Qy    234 cgttggaggccatgtgcangccgcgtacagctatgcgt 271
          |||||     || |||||||| | | ||| |   | | ||
Db  53561 CGTTGCGAGCTATGTGCAAGGCACGTGCCATTGTTGGT 53524
```

```
RESULT  12
ATAP21
LOCUS       ATAP21       206420 bp    DNA                PLN       30-JUL-1999
DEFINITION  Arabidopsis thaliana DNA chromosome 4, ESSA I AP2 contig fragment
            No. 1.
ACCESSION   Z99707
VERSION     Z99707.1  GI:4376087
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 206420)
  AUTHORS   Bevan,M., Terryn,N., Vos,P., Heijnen,L., Mewes,H.W., Mayer,K.F.X.
            and Schueller,C.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 206420)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUL-1999) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            schuelle@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
```

```
                    Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
                    E-mail: michael.bevan@bbsrc.ac.uk
COMMENT             On Mar 7, 1999 this sequence version replaced gi:4006849.
                    Information on performance of analysis and a more detailed
                    annotation of this entry and other sequences of chromosomes 3, 4
                    and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
                    this fragment has an overlap with ATAP22 at the 3' end.
FEATURES            Location/Qualifiers
     source         1. .206420
                    /organism="Arabidopsis thaliana"
                    /variety="Columbia"
                    /db_xref="taxon:3702"
                    /chromosome="4"
     source         1. .116845
                    /organism="Arabidopsis thaliana"
                    /db_xref="taxon:3702"
                    /clone="BAC TAMU8H13"
     exon           3. .560
                    /gene="C7A10.10"
                    /number=1
     gene           3. .560
                    /gene="C7A10.10"
     CDS            <3. .560
                    /gene="C7A10.10"
                    /note="similarity to cytochrome P450, Nicotiana tabacum,
                    PATX:E1188611
                    Contains Cytochrome P450 cysteine heme-iron ligand
                    signature [FGLGRRACPG]"
                    /codon_start=1
                    /product="cytochrome like protein"
                    /protein_id="CAB16768.1"
                    /db_xref="GI:4006850"
                    /translation="SLLNNPEVLNKARDEIDRMIGVDRLLEESDIPNLPYLQNIVSET
                    LRLYPAAPMLLPHVASKDCKVGGYDMPRGTMLLTNAWAIHRDPLLWDDPTSFKPERFE
                    KEGEAKKLMPFGLGRRACPGSGLAQRLVTLSLGSLIQCFEWERIGEEEVDMTEGPGLT
                    MPKARPLEAMCRARDFVGKILPDSS"
     gene           978. .2731
                    /gene="C7A10.20"
     CDS            join(978. .1475,1630. .2016,2114. .2731)
                    /gene="C7A10.20"
                    /note="similarity to cytochrome P450 Glycyrrhiza echinata,
                    PATX:D1023287
                    Contains Cytochrome P450 cysteine heme-iron ligand
                    signatures [FGLGRRACPG] and 474:[FGLGRRACPG]"
                    /codon_start=1
                    /product="cytochrome P450-like protein"
                    /protein_id="CAB16769.1"
                    /db_xref="GI:4006851"
                    /db_xref="SPTREMBL:O23154"
                    /translation="MEGQTLIFTFLFISLSLTFIIGRIKRRPNLPPSPSWALPVIGHL
                    RLLKPPLHRVFLSVSESLGDAPIISLRLGNRLVFVVSSHSLAEECFTKNDVVLANRFN
                    SLASKHISYGCTTVVTASYGDHWRNLRRIGAVEIFSAHRLNSFSSIRRDEIHRLIACL
                    SRNSSLEFTKVEMKSMFSNLTFNNIIRMLAGKCYYGDGAEDDPEAKRVRELIAEGMGC
                    FGAGNTADYLPILTWITGSEKRIKKIASRLDEFLQGLVDERREGKEKRQNTMVDHLLC
                    LQETQPEYYTDNIIKGIMLSLILAGTDTSAVTLEWTLSALLNHPQILSKARDEIDNKV
                    GLNRLVEESDLSHLPYLQNIVSESLRLYPASPLLVPHVASEDCKVGGYHMPRGTMLLT
                    NAWAIHRDPKIWDDPTSFKPERFEKEGEAQKLLGFGLGRRACPGSGLAQRLASLTIGS
```

```
                     LIQCFEWERIGEEEVDMTEGGGGVIMPKAIPLVAMCKARPVVGKILNESA"
exon                 978. .1475
                     /gene="C7A10.20"
                     /number=1
intron               1476. .1629
                     /gene="C7A10.20"
                     /number=1
exon                 1630. .2016
                     /gene="C7A10.20"
                     /number=2
intron               2017. .2113
                     /gene="C7A10.20"
                     /number=2
exon                 2114. .2731
                     /gene="C7A10.20"
                     /number=3
exon                 3115. .3615
                     /gene="C7A10.30"
                     /number=1
gene                 3115. .5137
                     /gene="C7A10.30"
CDS                  join(3115. .3615,4090. .4464,4535. .5137)
                     /gene="C7A10.30"
                     /note="strong similarity to cytochrome P450, Glycyrrhiza
                     echinata, PATX:D1023287
                     Contains Cytochrome P450 cysteine heme-iron ligand
                     signature 431:[FGMGRRACPG]
                     contains EST gb:T43640, Aa395149, T42716, T41670"
                     /codon_start=1
                     /product="cytochrome P450-like protein"
                     /protein_id="CAB16753.1"
                     /db_xref="GI:2464850"
                     /db_xref="SPTREMBL:O23155"
                     /translation="MDLNQILILSFLSLFTLAIFLLTRSKRKLNLPPSPAISLPVIGH
                     LHLLKPPLHRTFLSLSKSIGNAPVFHLRLGNRLVYVISSRSIAEECFTKNDVVLANRP
                     KFTISKHLGYNATYLLSASYGDHWRNLRRIAAVEIFSTHRLNSFLYIRKDEIRRLISH
                     LSRDSLHGFVEVEMKTLLTNLASNTTIRMLAGKRYFGEDNDDAKLVKNLVSEAVTSAG
                     AGNPIDYLSILRWVSSYEKRIKNLGNRFDTFLQKLVDEKRAEKEKGETMIDHLLALQD
                     IQPDYYTDVIIKGIILTLIIAGTDTSSVTLEWAMSNLLNHPEILKKARMEIDEKVGLD
                     RLVDESDIVNLSYLQSIVLETLRMYPAVPLLLPHLSSEDCKVGGYDIPSGTMVLTNAW
                     AMHRDPEVWEDPEIFKPERFEKEGEAEKLISFGMGRRACPGAGLAHRLINQALGSLVQ
                     CFEWERVGEDFVDMTEDKGATLPKAIPLRAMCKARSIVDKLI"
intron               3616. .4089
                     /gene="C7A10.30"
                     /number=1
exon                 4090. .4464
                     /gene="C7A10.30"
                     /number=2
intron               4465. .4534
                     /gene="C7A10.30"
                     /number=2
exon                 4535. .5137
                     /gene="C7A10.30"
                     /number=3
gene                 5994. .7942
                     /gene="C7A10.40"
CDS                  join(5994. .6494,6880. .7263,7340. .7942)
```

```
                        /gene="C7A10.40"
                        /note="similarity to cytochrome P450 monooxygenase,
                        Arabidopsis thaliana, PATchX:D1029478
                        Contains ATP/GTP-binding site motif A (P-loop) [ARAIVGKT],
                        Cytochrome P450 cysteine heme-iron ligand signature
                        [FGMGRRACPG]"
                        /codon_start=1
                        /product="cytochrome P450-like protein"
                        /protein_id="CAB16770.1"
                        /db_xref="GI:4376088"
                        /db_xref="SPTREMBL:O23156"
                        /translation="MDLTQILLLSFLFLTISIKLLLTKSNRKPNLPPSPAYPLPVIGH
                        LHLLKQPVHRTFHSISKSLGNAPIFHLRLGNRLVYVISSHSIAEECFTKNDVVLANRP
                        DIIMAKHVGYNFTNMIAASYGDHWRNLRRIAAVEIFSSHRISTFSSIRKDEIRRLITH
                        LSRDSLHGFVEVELKSLLTNLAFNNIIMMVAGKRYYGTGTEDNDEAKLVRELIAEIMA
                        GAGSGNLADYLPSINWVTNFENQTKILGNRLDRVLQKLVDEKRAEKEKGQTLIDHLLS
                        FQETEPEYYTDVIIKGIILALVLAGTDTSSVTLEWAMSNLLNHPEILEKARAEIDDKI
                        GSDRLVEESDIVNLHYLQNIVSETLRLYPAVPLLLPHFSSDECKVAGYDMPRRTLLLT
                        NVWAMHRDPGLWEEPERFKPERFEKEGEARKLMPFGMGRRACPGAELGKRLVSLALGC
                        LIQSFEWERVGAELVDMTEGEGITMPKATPLRAMCKARAIVGKTI"
exon                    5994..6494
                        /gene="C7A10.40"
                        /number=1
intron                  6495..6879
                        /gene="C7A10.40"
                        /number=1
exon                    6880..7263
                        /gene="C7A10.40"
                        /number=2
intron                  7264..7339
                        /gene="C7A10.40"
                        /number=2
exon                    7340..7942
                        /gene="C7A10.40"
                        /number=3
exon                    8851..9360
                        /gene="C7A10.50"
                        /number=1
gene                    8851..11532
                        /gene="C7A10.50"
CDS                     join(8851..9360,9980..10369,10876..11532)
                        /gene="C7A10.50"
                        /note="similarity to cytochrome P450 Glycyrrhiza echinata,
                        PATCHX:D102328
                        Contains Cytochrome P450 cysteine heme-iron ligand
                        signature [FGLGRRACPG]
                        contains EST gb:F13573, F13574"
                        /codon_start=1
                        /product="cytochrome P450-like protein"
                        /protein_id="CAB16771.1"
                        /db_xref="GI:4006853"
                        /translation="MDCILLILTTLVAIFIVKIVLLVTKPNKNLPPSPNICFPIIGHL
                        HLLKKPLLHRTLSHLSHSLGPVFSLRLGSRLAVIISSPTAAEECFLTKNDIVLANRPR
                        FIMGKYVAYDYTSMVTAPYGDHWRNLRRITALEVFSTNRLNASAEIRHDEVKMLLQKL
                        HDLSVERPAKVELRQLLTGLTLNVIMRMMTGKRFFEEDDGGKAGISLEFRELVAEILE
                        LSAADNPADFLPALRWFDYKGLVKRAKRIGERMDSLLQGFLDEHRANKDRLEFKNTMI
                        AHLLDSQEKEPHNYSDQTIKGLILMMVVGGTDTSALTVEWAMSNLLNHPQILETTRQN
```

```
                    IDTQMETSSSRRLLKEEDLVNMNYLKNVVSETLRLYPVAPLMVPHVPSSDCVIGGFNV
                    PRDTIVLVNLWAIHRDPSVWDDPTSFKPERFEGSDQFGHYNGKMMPFGLGRRACPGLS
                    LANRVVGLLLGSMIQCFEWESGSGGQVDMTEGPGLSLPKAEPLVVTCRTREMASELLF
                    FGSEPSNKNV"
     intron         9361. .9979
                    /gene="C7A10.50"
                    /number=1

  Query Match              13.7%;   Score 59.2;  DB 8;  Length 206420;
  Best Local Similarity    54.1%;   Pred. No. 8.7e-05;
  Matches  118;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;


Qy      54  aggggcccctgctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcg  113
            ||| |     ||| || || || ||||||||| |·| |     ||| || || |    ||| |
Db    7713  AGGCTCGAAAGCTAATGCCGTTTGGGATGGGACGACGAGCTTGTCCTGGAGCTGAGCTTG  7772


Qy     114  cgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacgg  173
            |· |||| || | || ||| |    || | | ||||| |||||| |
Db    7773  GGAAGCGGTTAGTGAGCCTTGCTCTTGGGTGCTTGATTCAGTCTTTCGAGTGGGAGAGAG  7832


Qy     174  ttgatggagctcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcc  233
            ·||| || || | ||      | || || ||| | || ||||| || |
Db    7833  TTGGTGCAGAACTTGTGGACATGACTGAAGGCGAAGGGATCACTATGCCTAAAGCTACTC  7892


Qy     234  cgttggaggccatgtgcangccgcgtacagctatgcgt  271
            ||||·| || ||||||||| | · ||| |   | | ||
Db    7893  CGTTGCGAGCTATGTGCAAGGCACGTGCCATTGTTGGT  7930



RESULT  13
AY039844
LOCUS       AY039844      1656 bp     mRNA            PLN        24-JUN-2001
DEFINITION  Arabidopsis thaliana AT4g37430/F6G17_80 mRNA, complete cds.
ACCESSION   AY039844
VERSION     AY039844.1  GI:14532439
KEYWORDS    FLI_CDNA.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1656)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
            Bowser,L., Carninci,P., Dale,J.M., Gibson,H.A., Goldsmith,A.D.,
            Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A.,
            Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X.,
            Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
            Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,
            Tang,C.C., Toriumi,M., Yamada,K., Yu,G., Yu,S., Shinozaki,K.,
            Davis,R.W., Theologis,A. and Ecker,J.R.
  TITLE     Arabidopsis cDNA clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1656)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
            Bowser,L., Carninci,P., Dale,J.M., Gibson,H.A., Goldsmith,A.D.,
            Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A.,
            Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X.,
```

Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
  TITLE   Direct Submission
  JOURNAL   Submitted (06-JUN-2001) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
  COMMENT   RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
            Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J. Bowser,L.,
            Dale,J.M., Gibson,H.A., Goldsmith,A.D., Jiang,P.X., Jones,T.,
            Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
            Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L.,
            Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yu,G., Yu,S.,
            Davis,R.W., Theologis,A., and Ecker,J.R.

            Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES   Location/Qualifiers
     source   1. .1656
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /clone="RAFL05-02-B21(R12849)"
                     /note="ecotype: Columbia"
     5'UTR   1. .59
     CDS   60. .1562
                     /note="cytochrome P450 monooxygenase (CYP91A2)"
                     /codon_start=1
                     /product="AT4g37430/F6G17_80"
                     /protein_id="AAK63948.1"
                     /db_xref="GI:14532440"
                     /translation="MLYFILLPLLFLVISYKFLYSKTQRFNLPPGPPSRPFVGHLHLM
                     KPPIHRLLQRYSNQYGPIFSLRFGSRRVVVITSPSLAQESFTGQNDIVLSSRPLQLTA
                     KYVAYNHTTVGTAPYGDHWRNLRRICSQEILSSHRLINFQHIRKDEILRMLTRLSRYT
                     QTSNESNDFTHIELEPLLSDLTFNNIVRMVTGKRYYGDDVNNKEEAELFKKLVYDIAM
                     YSGANHSADYLPILKLFGNKFEKEVKAIGKSMDDILQRLLDECRRDKEGNTMVNHLIS
                     LQQQQPEYYTDVIIKGLMMSMMLAGTETSAVTLEWAMANLLRNPEVLEKARSEIDEKI
                     GKDRLIDESDIAVLPYLQNVVSETFRLFPVAPFLIPRSPTDDMKIGGYDVPRDTIVMV
                     NAWAIHRDPEIWEEPEKFNPDRYNDGCGSDYYVYKLMPFGNGRRTCPGAGLGQRIVTL
                     ALGSLIQCFEWENVKGEEMDMSESTGLGMRKMDPLRAMCRPRPIMSKLLL"
     3'UTR   1563. .1656
BASE COUNT   503 a   410 c   348 g   395 t
ORIGIN

  Query Match   13.7%; Score 59; DB 8; Length 1656;
  Best Local Similarity   52.7%; Pred. No. 0.00014;

```
      Matches  125;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy       32 gatggctccggcggcaaggccaaggggcccctgctgatccctttcgggatggggcggccc 91
            || ||  | |||  || |    | | |     |  |||||| || || |||| || |||
Db·    1308 GACGGATGCGGAAGCGATTACTATGTTTACAAGCTGATGCCGTTTGGGAATGGCCGGAGA 1367

Qy       92 aattgccccgggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctc 151
            | |||  |||||| |       | | | | |||  ||       ||| ||| || || | |
Db     1368 ACTTGTCCCGGCGCCGGATTAGGTCAGAGGATTGTGACATTGGCGCTTGGATCGTTGATT 1427

Qy      152 aattgcttcgactgggacacggttgatggagctcaggtttgacatgaagctancggcggg 211
             | |||||| || |||||| |  ||  | || |  || |        |  |      |||
Db     1428 CAATGCTTTGAATGGGAGAATGTGAAAGGGGAAGAGATGGATATGTCTGAGAGTACTGGG 1487

Qy      212 ctgaccatgccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatg 268
            ||    ||||     |  |  || ||  ||||||||||| | |||  |  |    ||||
Db     1488 TTGGGTATGCGTAAGATGGATCCTTTACGGGCCATGTGTAGGCCTAGGCCCATTATG 1544


RESULT   14
ATF6G17
LOCUS       ATF6G17      101009 bp     DNA             PLN        03-MAR-1999
DEFINITION  Arabidopsis thaliana DNA chromosome 4, BAC clone  F6G17 (ESSA
            project).
ACCESSION   AL035601
VERSION     AL035601.1  GI:4468801
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 101009)
  AUTHORS   Bevan,M., Koetter,P., Hempel,S., Entian,K.-D., Bancroft,I.,
            Mewes,H.W., Mayer,K.F.X. and Schueller,C.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 101009)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAR-1999) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            schuelle@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosome 4 can be
            viewed at: http://websvr.mips.biochem.mpg.de/proj/thal/.
FEATURES             Location/Qualifiers
     source          1. .101009
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     misc_feature    1. .18845
                     /note="position 1-18845 overlap to EMBL accession Z99707;
                     please refer to this entry for analysis and annotation"
```

```
gene            complement(join(18286. .18900,19005. .19391,19523. .20020))
                /gene="F6G17.10"
gene            18286. .20020
                /gene="F6G17.10"
CDS             complement(join(18286. .18900,19005. .19391,19523. .20020))
                /gene="F6G17.10"
                /note="similarity to cytochrome P450 monooxygenase,
                Arabidopsis thaliana, D78606
                Contains Cytochrome P450 cysteine heme-iron ligand
                signature[FGLGRRACPG]"
                /codon_start=1
                /product="cytochrome p450-like protein"
                /protein_id="CAB38203.1"
                /db_xref="GI:4468802"
                /translation="MEALMLIFTFCFIVLSLIFLIGRIKRKLNLPPSPAWALPVIGHL
                RLLKPPLHRVFLSVSQSLGDAPIISLRLGNRLLFVVSSHSIAEECFTKNDVILANRQT
                TISTKHISYGNSTVVSASYSEHWRNLRRIGALEIFSAHRLNSFSSIRRDEIRRLIGRL
                LRNSSYGFTKVEMKSMFSDLTFNNIIRMLAGKCYYGDGKEDDPEAKRVRTLIAEAMSS
                SGPGNAADYIPILTWITYSETRIKKLAGRLDEFLQGLVDEKREGKEKKENTMVDHLLC
                LQETQPEYYMDRIIKGTMLSLIAGGTDTTAVTLEWALSSLLNNPEVLNKARDEIDRMI
                GVDRLLEESDIPNLPYLQNIVSETLRLYPAAPMLLPHVASKDCKVGGYDMPRGTMLLT
                NAWAIHRDPLLWDDPTSFKPERFEKEGEAKKLMPFGLGRRACPGSGLAQRLVTLSLGS
                LIQCFEWERIGEEEVDMTEGPGLTMPKARPLEAMCRARDFVGKILPDSS"
exon            complement(18286. .18900)
                /gene="F6G17.10"
                /number=1
intron          complement(18901. .19004)
                /number=1
exon            complement(19005. .19391)
                /gene="F6G17.10"
                /number=2
intron          complement(19392. .19522)
                /number=2
exon            complement(19523. .20020)
                /gene="F6G17.10"
                /number=3
gene            21133. .22840
                /gene="F6G17.20"
exon            complement(21133. .21744)
                /gene="F6G17.20"
                /number=1
gene            complement(join(21133. .21744,21842. .22225,22343. .22840))
                /gene="F6G17.20"
CDS             complement(join(21133. .21744,21842. .22225,22343. .22840))
                /gene="F6G17.20"
                /note="similarity to cytochrome P450, Glycyrrhiza
                echinata, AB001379
                Contains Cytochrome P450 cysteine heme-iron ligand
                signature [FGLGRRACPG]
                contains EST gb:AA586064, H76015, T41596, N38867"
                /codon_start=1
                /product="cytochrome P450-like protein"
                /protein_id="CAB38204.1"
                /db_xref="GI:4468803"
                /translation="METKTLIFSILFVVLSLIYLIGKLKRKPNLPPSPAWSLPVIGHL
                RLLKPPIHRTFLSLSQSLNNAPIFSLRLGNRLVFVNSSHSIAEECFTKNDVVLANRPN
                FILAKHVAYDYTTMIAASYGDHWRNLRRIGSVEIFSNHRLNSFLSIRKDEIRRLVFRL
```

```
                         SRNFSQEFVKVDMKSMLSDLTFNNILRMVAGKRYYGDGVEDDPEAKRVRQLIADVVAC
                         AGAGNAVDYLPVLRLVSDYETRVKKLAGRLDEFLQGLVDEKREAKEKGNTMIDHLLTL
                         QESQPDYFTDRIIKGNMLALILAGTDTSAVTLEWALSNVLNHPDVLNKARDEIDRKIG
                         LDRLMDESDISNLPYLQNIVSETLRLYPAAPMLLPHVASEDCKVAGYDMPRGTILLTN
                         VWAIHRDPQLWDDPMSFKPERFEKEGEAQKLMPFGLGRRACPGSGLAHRLINLTLGSL
                         IQCLEWEKIGEEVDMSEGKGVTMPKAKPLEAMCRARPSVVKIFNESV"
intron                   complement(21745. .21841)
                         /number=1
exon                     complement(21842. .22225)
                         /gene="F6G17.20"
                         /number=2
intron                   complement(22226. .22342)
                         /number=2
exon                     complement(22343. .22840)
                         /gene="F6G17.20"
                         /number=3
gene                     23202. .25100
                         /gene="F6G17.30"
CDS                      complement(23202. .25100)
                         /gene="F6G17.30"
                         /note="similarity to various predicted proteins,
                         Arabidopsis thaliana
                         Contains Cytochrome c family heme-binding site signature
                         [CSDCHT]"
                         /codon_start=1
                         /product="putative protein"
                         /protein_id="CAB38205.1"
                         /db_xref="GI:4468804"
                         /translation="MASSPLLATSLPQNQLSTTATARFRLPPPEKLAVLIDKSQSVDE
                         VLQIHAAILRHNLLLHPRYPVLNLKLHRAYASHGKIRHSLALFHQTIDPDLFLFTAAI
                         NTASINGLKDQAFLLYVQLLSSEINPNEFTFSSLLKSCSTKSGKLIHTHVLKFGLGID
                         PYVATGLVDVYAKGGDVVSAQKVFDRMPERSLVSSTAMITCYAKQGNVEAARALFDSM
                         CERDIVSWNVMIDGYAQHGFPNDALMLFQKLLAEGKPKPDEITVVAALSACSQIGALE
                         TGRWIHVFVKSSRIRLNVKVCTGLIDMYSKCGSLEEAVLVFNDTPRKDIVAWNAMIAG
                         YAMHGYSQDALRLFNEMQGITGLQPTDITFIGTLQACAHAGLVNEGIRIFESMGQEYG
                         IKPKIEHYGCLVSLLGRAGQLKRAYETIKNMNMDADSVLWSSVLGSCKLHGDFVLGKE
                         IAEYLIGLNIKNSGIYVLLSNIYASVGDYEGVAKVRNLMKEKGIVKEPGISTIEIENK
                         VHEFRAGDREHSKSKEIYTMLRKISERIKSHGYVPNTNTVLQDLEETEKEQSLQVHSE
                         RLAIAYGLISTKPGSPLKIFKNLRVCSDCHTVTKLISKITGRKIVMRDRNRFHHFTDG
                         SCSCGDFW"
exon                     complement(23202. .25100)
                         /gene="F6G17.30"
                         /number=1
gene                     complement(23202. .25100)
                         /gene="F6G17.30"
gene                     30884. .32930
                         /gene="F6G17.40"
exon                     30884. .31206
                         /gene="F6G17.40"
                         /number=1
CDS                      join(30884. .31206,31359. .31460,31544. .32930)
                         /gene="F6G17.40"
                         /note="similarity to hypothetical protein, Glycine max,
                         PIR2:S17433
                         contains EST gb:T44614, AAA712644, AA395496"
                         /codon_start=1
                         /product="auxin-responsive GH3-like protein"
```

```
                          /protein_id="CAB38206.1"
                          /db_xref="GI:4468805"
                          /translation="MAVDSPLQSRMVSATTSEKDVKALKFIEEMTRNPDSVQEKVLGE
                          ILTRNSNTEYLKRFDLDGVVDRKTFKSKVPVVTYEDLKPEIQRISNGDCSPILSSHPI
                          TEFLTSSGTSAGERKLMPTIEEDLDRRQLLYSLLMPVMNLYVPGLDKGKGLYFLFVKS
                          ESKTSGGLPARPVLTSYYKSDHFKRRPYDPYNVYTSPNEAILCSDSSQSMYAQMLCGL
                          LMRHEVLRLGAVFASGLLRAISFLQNNWKELARDISTGTLSSRIFDPAIKNRMSKILT
                          KPDQELAEFLVGVCSQENWEGIITKIWPNTKYLDVIVTGAMAQYIPTLEYYSGGLPMA
                          CTMYASSESYFGINLKPMCKPSEVSYTIMPNMAYFEFLPHNHDGDGAAEASLDETSLV
                          ELANVEVGKEYELVITTYAGLYRYRVGDILRVTGFHNSAPQFKFIRRKNVLLSVESDK
                          TDEAELQKAVENASRLFAEQGTRVIEYTSYAETKTIPGHYVIYWELLGRDQSNALMSE
                          EVMAKCCLEMEESLNSVYRQSRVADKSIGPLEIRVVRNGTFEELMDYAISRGASINQY
                          KVPRCVSFTPIMELLDSRVVSAHFSPSLPHWSPERRR"
     intron            31207. .31358
                          /gene="F6G17.40"
                          /number=1
     exon              31359. .31460
                          /gene="F6G17.40"
                          /number=2
     intron            31461. .31543
                          /gene="F6G17.40"
                          /number=2
     exon              31544. .32930
                          /gene="F6G17.40"
                          /number=3
     gene              35807. .37359
                          /gene="F6G17.50"
     CDS               join(35807. .36127,36230. .36289,36724. .37359)
                          /gene="F6G17.50"
                          /note="strong similarity to cytochrome P450 monooxygenase
                          CYP91A2, Arabidopsis thaliana, D78607
                          Contains Cytochrome P450 cysteine heme-iron ligand
                          signature [FGNGRRSCPG]
                          contains EST gb:AA712784"
                          /codon_start=1
                          /product="cytochrome P450 monooxygenase-like protein"
                          /protein_id="CAB38207.1"
                          /db_xref="GI:4468806"
                          /translation="MVTGKRYYGDEVHNEEEANVFKKLVADINDCSGARHPGDYLPFM
                          KMFGGSFEKKVKALAEAMDEILQRLLEECKRDKDGNTMVNHLLSLQQNEPEYYTDVTI
                          KGLMLIFCFFGQLQILWFTNIETGWGMMIAGTDTSAVTLEWAMSSLLNHPEALEKAKL
                          EIDEKIGQERLIDEPDIANLPYLQNIVSETFRLYPAAPLLVPRSPTEDIKVGGYDVPR
                          GTMVMVNAWAIHRDPELWNEPEKFKPERFNGGEGGGRGEDVHKLMPFGNGRRSCPGAG
                          LGQKIVTLALGSLIQCFDWQKVNGEAIDMTETPGMAMRKKIPLSALCQSRPIMSKLQA
                          HLKG"
     exon              35807. .36127
                          /gene="F6G17.50"
                          /number=1
     intron            36128. .36229
                          /gene="F6G17.50"


     Query Match              13.7%;   Score 59;  DB 8;  Length 101009;
     Best Local Similarity    52.7%;   Pred. No. 0.0001;
     Matches   125;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps      0;


Qy     32 gatggctccggcggcaaggccaaggggcccctgctgatcccctttcgggatggggcggccc 91
          || || | |||  || |   | | |      |  |||||| || || |||| || |||
```

```
Db   49737  GACGGATGCGGAAGCGATTACTATGTTTACAAGCTGATGCCGTTTGGGAATGGCCGGAGA 49796

Qy      92  aattgccccgggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctc 151
            |  ||| |||||| |       | |  | | ||| ||      ||| ||| | | || || |
Db   49797  ACTTGTCCCGGCGCCGGATTAGGTCAGAGGATTGTGACATTGGCGCTTGGATCGTTGATT 49856

Qy     152  aattgcttcgactgggacacggttgatggagctcaggtttgacatgaagctancggcggg 211
            '| |||||| || ||||| | || | ||| |   || |          .| |       ||| 
Db   49857  CAATGCTTTGAATGGGAGAATGTGAAAGGGGAAGAGATGGATATGTCTGAGAGTACTGGG 49916

Qy     212  ctgaccatgccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatg 268
            ||    ||||     |  |   || ·||  |||||||||| | |||  |  |    |||| 
Db   49917  TTGGGTATGCGTAAGATGGATCCTTTACGGGCCATGTGTAGGCCTAGGCCCATTATG 49973
```

RESULT 15
ATCHRIV87

```
LOCUS       ATCHRIV87   196339 bp     DNA             PLN        16-MAR-2000
DEFINITION  Arabidopsis thaliana DNA chromosome 4, contig fragment No. 87.
ACCESSION   AL161591
VERSION     AL161591.2  GI:7270703
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 42610 to 143618; 123423 to 196339)
  AUTHORS   Rose,M., Hempel,S., Entian,K.-D., Mewes,H.W., Lemcke,K. and
            Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196339)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
            this fragment has an overlap with ATCHRIV86 at the 5' end and an
            overlap with ATCHRIV88 at the 3' end.
FEATURES             Location/Qualifiers
     source          1. .196339
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     gene            6146. .7792
                     /gene="AT4g37210"
     exon            6146. .6474
                     /gene="AT4g37210"
                     /number=1
     CDS             join(6146. .6474,6702. .6919,7013. .7225,7313. .7403,
```

```
                          7489. .7792)
                          /gene="AT4g37210"
                          /note="intron number 1 is a special U12 intron
                          similarity to nuclear histone-binding protein N1/N2,
                          Xenopus laevis, PIR2:A25680"
                          /codon_start=1
                          /product="putative protein"
                          /protein_id="CAB80387.1"
                          /db_xref="GI:7270704"
                          /translation="MVEESASASEASVIQTLTEPATEIAQTLEPNLASIEATVESVVQ
                          GGTESTCNNDANNNNAADSAATEVCDEEREKTLEFAEELTEKGSVFLKENDFAEAVDC
                          FSRALEIRVAHYGELDAECINAYYRYGLALLAKAQAEADPLGNMPKKEGEVQQESSNG
                          ESLAPSVVSGDPERQGSSSGQEGSGGKDQGEDGEDCQDDDLSDADGDADEDESDLDMA
                          WKMLDIARVITDKQSTETMEKVDILCSLAEVSLEREDIESSLSDYKNALSILERLVEP
                          DSRRTAELNFRICICLETGCQPKEAIPYCQKALLICKARMERLSNEIKGASGSATSST
                          VSEIDEGIQQSSNVPYIDKSASDKEVEIGDLAGLAEDLEKKASKLNLSVH"
     intron               6475. .6701
                          /gene="AT4g37210"
                          /number=1
     exon                 6702. .6919
                          /gene="AT4g37210"
                          /number=2
     intron               6920. .7012
                          /gene="AT4g37210"
                          /number=2
     exon                 7013. .7225
                          /gene="AT4g37210"
                          /number=3
     intron               7226. .7312
                          /gene="AT4g37210"
                          /number=3
     exon                 7313. .7403
                          /gene="AT4g37210"
                          /number=4
     intron               7404. .7488
                          /gene="AT4g37210"
                          /number=4
     exon                 7489. .7792
                          /gene="AT4g37210"
                          /number=5
     gene                 8876. .9739
                          /gene="AT4g37220"
     CDS                  join(8876. .9051,9141. .9248,9330. .9400,9486. .9739)
                          /gene="AT4g37220"
                          /note="strong similarity to cold acclimation protein
                          WCOR413, Triticum aestivum, PATCHX:G1657855
                          Contains Prokaryotic membrane lipoprotein lipid attachment
                          site AA147-157
                          contains EST gb:AW033651.1, W43270, AA650647, AI996990.1,
                          AA728669, T42795, Z37671, AI100742, T42949, AA040998,
                          AA395771, AA657303, T41871, T45633"
                          /codon_start=1
                          /product="cold acclimation protein homolog"
                          /protein_id="CAB80388.1"
                          /db_xref="GI:7270705"
                          /translation="MGRGEFLAMKTEENAANLINSDMNEFVAAAKKLVKDVGMLGGVG
                          FGTSVLQWAASIFAIYLLILDRTNWKTKMLTTLLVPYIFFTLPSVIFQFFSGDFGKWI
```

```
                          ALIAIIVRLFFPKEFPEWLEIPVALILIVVVSPSLIAWTLRESWVGAVICLVIACYLF
                          HEHIKASGGFKNSFTQKNGISNTIGIVALLVYPVWTIFFHIF"
     exon                 8876. .9051
                          /gene="AT4g37220"
                          /number=1
     intron               9052. .9140
                          /gene="AT4g37220"
                          /number=1
     exon                 9141. .9248
                          /gene="AT4g37220"
                          /number=2
     intron               9249. .9329
                          /gene="AT4g37220"
                          /number=2
     exon                 9330. .9400
                          /gene="AT4g37220"
                          /number=3
     intron               9401. .9485
                          /gene="AT4g37220"
                          /number=3
     exon                 9486. .9739
                          /gene="AT4g37220"
                          /number=4
     gene                 10194. .11357
                          /gene="AT4g37230"
     exon                 10194. .10398
                          /gene="AT4g37230"
                          /number=1
     CDS                  join(10194. .10398,11131. .11357)
                          /gene="AT4g37230"
                          /note="possible frameshift in DNA sequence at pos.
                          51188-51200, 5' part of gene couldn't be reconstructed,
                          possible pseudogene, no ATG
                          strong similarity to photosystem II oxygen-evolving
                          complex protein 1, spinach, PIR2:A23626
                          contains EST gb:Z34685"
                          /codon_start=1
                          /product="photosystem II oxygen-evolving complex like
                          protein (partial)"
                          /protein_id="CAB80389.1"
                          /db_xref="GI:7270706"
                          /translation="MTSLTYTLDEIEGPFEVDYAAVTVHNFLVGSVYRSCSRSSSSWH
                          RRGYFGPKSIPSAFTQGHVGNKSDQYQGYDNAVALPARGNNEELAKENNKITLSVTKS
                          NPESGEVIGAFESIQPSDTDLGATTPKDVKIQGIWYCQLDE"
     intron               10399. .11130
                          /gene="AT4g37230"
                          /number=1
     exon                 11131. .11357
                          /gene="AT4g37230"
                          /number=2
     exon                 18388. .18822
                          /gene="AT4g37240"
                          /number=1
     gene                 18388. .18822
                          /gene="AT4g37240"
     CDS                  18388. .18822
                          /gene="AT4g37240"
```

```
                           /note="contains EST gb:Z18456"
                           /codon_start=1
                           /product="putative protein"
                           /protein_id="CAB80390.1"
                           /db_xref="GI:7270707"
                           /translation="MEFANPVKVGYVLLKYPMCFICNSDDMDFDDAVAAISADEELQL
                    GQIYFALPLCWLRQPLKAEEMAALAVKASSALMRGGGGGCRRKCVEPIVSDKLRMRVG
                    SGDDTVGSGSGRRKVRNGDGGGSVSSSRRRKCYAAELSTIDE"
          gene             21559. .23955
                           /gene="AT4g37250"
          exon             complement(21559. .22253)
                           /gene="AT4g37250"
                           /number=1
          gene             complement(join(21559. .22253,22350. .23955))
                           /gene="AT4g37250"
          CDS              complement(join(21559. .22253,22350. .23955))
                           /gene="AT4g37250"
                           /note="similarity to protein kinase TMK1, receptor type
                    precursor - Arabidopsis thaliana, PIR1:JQ1674
                    contains EST gb:H76836"
                           /codon_start=1
                           /product="receptor kinase-like protein"
                           /protein_id="CAB80391.1"
                           /db_xref="GI:7270708"
                           /translation="MELISVIFFFFCSVLSSSALNSDGLVLMKFKSSVLVDPLSLLQT
                    WNYKHESPCSWRGISCNNDSKVLTLSLPNSQLLGSIPSDLGSLLTLQSLDLSNNSFNG
                    PLPVSFFNARELRFLDLSSNMISGEIPSAIGDLHNLLTLNLSDNALAGKLPTNLASLR
                    NLTVVSLENNYFSGEIPGGWRVVEFLDLSSNLINGSLPPDFGGYSLQYLNVSFNQISG
                    EIPPEIGVNFPRNVTVDLSFNNLTGPIPDSPVFLNQESNFFSGNPGLCGEPTRNPCLI
                    PSSPSIVSEADVPTSTPAIAAIPNTIGSNPVTDPNSQQTDPNPRTGLRPGVIIGIVVG
                    DIAGIGILAVIFLYIYRCKKNKIVDNNNNDKQRTETDTITLSTFSSSSSSPEESRRFR
                    KWSCLRKDPETTPSEEEDEDDEDEESGYNANQRSGDNKLVTVDGEKEMEIETLLKASA
                    YILGATGSSIMYKAVLEDGRVFAVRRLGENGLSQRRFKDFEPHIRAIGKLVHPNLVRL
                    CGFYWGTDEKLVIYDFVPNGSLVNPRYRKGGGSSSPYHLPWETRLKIAKGIARGLAYL
                    HEKKHVHGNLKPSNILLGHDMEPKIGDFGLERLLTGETSYIRAGGSSRIFSSKRYTTS
                    SREFSSIGPTPSPSPSSVGAMSPYCAPESFRSLKPSPKWDVYGFGVILLELLTGKIVS
                    VEEIVLGNGLTVEDGHRAVRMADVAIRGELDGKQEFLLDCFKLGYSCASPVPQKRPTM
                    KESLAVLERFHPNSSVIKSSSFHYGH"
          intron           complement(22254. .22349)
                           /number=1
          exon             complement(22350. .23955)
                           /gene="AT4g37250"
                           /number=2
          gene             34372. .35334
                           /gene="AT4g37260"
          CDS              34372. .35334
```

```
    Query Match            13.7%;  Score 59;  DB 8;  Length 196339;
    Best Local Similarity  52.7%;  Pred. No. 9.8e-05;
    Matches  125;  Conservative   0;  Mismatches 112;  Indels   0;  Gaps    0;

Qy       32 gatggctccggcggcaaggccaaggggcccctgctgatccctttcgggatggggcggccc 91
            || || | ||| || |   | | |    |  |||||| || || |||| || |||
Db    92346 GACGGATGCGGAAGCGATTACTATGTTTACAAGCTGATGCCGTTTGGGAATGGCCGGAGA 92405

Qy       92 aattgccccgggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctc 151
            | ||| |||||| |    | |  | | ||| ||    ||| ||| | | || || |
```

```
Db   92406 ACTTGTCCCGGCGCCGGATTAGGTCAGAGGATTGTGACATTGGCGCTTGGATCGTTGATT 92465

Qy     152 aattgcttcgactgggacacggttgatggagctcaggtttgacatgaagctancggcggg 211
               |  ||||||  ||  ||||||  |   ||   |  || |    || |          |   |      | | |
Db   92466 CAATGCTTTGAATGGGAGAATGTGAAAGGGGAAGAGATGGATATGTCTGAGAGTACTGGG 92525

Qy     212 ctgaccatgccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatg 268
               ||    ||||     |     |  || ||    ||||||||||| | |||  |   |    ||||
Db   92526 TTGGGTATGCGTAAGATGGATCCTTTACGGGCCATGTGTAGGCCTAGGCCCATTATG 92582
```


Search completed: February 7, 2002, 11:08:54
Job time: 10060 sec

                         GenCore version 4.5
                Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2002, 11:00:03 ; Search time 428.31 Seconds
                                             (without alignments)
                                             864.711 Million cell updates/sec

Title: US-09-394-745-6514
Perfect score: 432
Sequence: 1 gtccagcagctcggacttac..........attttctttttttttcttgg 432

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_1101:*
                1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
                2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
                3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
                4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
                5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
                6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
                7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
                8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
                9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
                10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
                11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
                12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
                13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*

```
        14:  /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
        15:  /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
        16:  /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
        17:  /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
      ᴗ 18:  /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
        19:  /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
        20:  /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
        21:  /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
        22:  /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	59.6	13.8	622	19	AAV23837	Plant C4H enzyme D
2	59.6	13.8	622	20	AAZ06838	Eucalyptus cinnama
3	59.6	13.8	622	21	AAA67924	Eucalyptus grandis
4	59	13.7	1655	21	AAC47389	Arabidopsis thalia
5	58	13.4	1656	21	AAC37476	Arabidopsis thalia
6	56.6	13.1	1674	20	AAX58406	Jerusalem artichok
7	53.4	12.4	1859	21	AAA29326	Soybean isoflavone
8	51.6	11.9	1817	14	AAQ50511	Bx1 gene. Zea may
9	49.4	11.4	1847	21	AAZ50024	Maize cytochrome p
10	49.4	11.4	1848	21	AAZ87320	Maize cytochrome P
11	48.8	11.3	1545	21	AAC42545	Arabidopsis thalia
12	48.8	11.3	1576	21	AAC47416	Arabidopsis thalia
13	48.8	11.3	1578	21	AAC35968	Arabidopsis thalia
14	46.8	10.8	1497	21	AAC50087	Arabidopsis thalia
15	43.8	10.1	1519	21	AAC47053	Arabidopsis thalia
16	43.6	10.1	1494	21	AAC43264	Arabidopsis thalia
17	43.6	10.1	1654	21	AAC50097	Arabidopsis thalia
18	42	9.7	1506	20	AAX26871	Codon modified cDN
19	42	9.7	1506	20	AAX26872	Codon modified cDN
20	42	9.7	1506	20	AAX26873	Codon modified cDN
21	42	9.7	1653	21	AAC41374	Arabidopsis thalia
22	42	9.7	2261	20	AAX26865	Cytochrome P450 cD
23	41.8	9.7	1664	21	AAC47894	Arabidopsis thalia
24	41.8	9.7	1665	21	AAC50095	Arabidopsis thalia
25	40.2	9.3	1745	18	AAT94656	Carnation flavonoi
26	39.8	9.2	447	21	AAA68128	Eucalyptus grandis
27	39.8	9.2	868	22	AAH87698	Peppermint plant o
28	39.8	9.2	1698	21	AAA29327	Soybean isoflavone
29	39.6	9.2	1866	21	AAA68010	Eucalyptus grandis
30	39.2	9.1	508	21	AAC57008	Pinus radiata tran
31	39	9.0	755	20	AAX25212	Maize ferulate-5-h
32	39	9.0	1545	22	AAS05178	Taxus cuspidata ox
33	39	9.0	2086	22	AAD05733	Maize ferulate-5-h
34	38.4	8.9	1696	21	AAC81819	Gerbera flavone sy
35	38.4	8.9	1762	20	AAX06560	(-)-limonene-6-hyd
36	38.4	8.9	1762	21	AAZ52307	Mint carveol synth
37	38.4	8.9	1762	22	AAF76606	Spearmint (-)-limo
38	38.4	8.9	1762	22	AAF76662	Computer-generated

39	38.4	8.9	1762	22	AAF76663	Computer-generated
40	38	8.8	1730	21	AAA87997	Torenia hybrida fl
41	37.6	8.7	1824	18	AAT94663	Morning glory flav
42	37	8.6	1929	19	AAV57472	Sorghum bicolor (L
43	36.4	8.4	375	21	AAH51157	Human CYP1A2 relat
44	36.2	8.4	1659	21	AAC50983	Arabidopsis thalia
45	36.2	8.4	1662	21	AAC39341	Arabidopsis thalia

ALIGNMENTS

RESULT   1
AAV23837
ID   AAV23837 standard; DNA; 622 BP.
XX
AC   AAV23837;
XX
DT   31-JUL-1998   (first entry)
XX
DE   Plant C4H enzyme DNA sequence.
XX
KW   Lignin biosynthetic pathway; eucalyptus; pine; transgenic plant;
KW   lignin content; tree processing; cellulose fibre; ss.
XX
OS   Eucalyptus grandis.
XX
PN   WO9811205-A2.
XX
PD   19-MAR-1998.
XX
PF   10-SEP-1997;   97WO-NZ00112.
XX
PR   11-SEP-1996;   96US-0713000.
XX
PA   (FLET-) FLETCHER CHALLENGE FORESTS LTD.
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Bloksberg LN, Grierson AW, Havukkala IJ;
XX
DR   WPI; 1998-207374/18.
XX
PT   Sequences useful for modification of plant lignin content or
PT   structure - from Eucalyptus grandis (eucalyptus) and Pinus radiata
PT   (pine) are associated with lignin biosynthesis pathway, useful e.g.
PT   in paper industry
XX
PS   Claim 1; Page 35; 82pp; English.
XX
CC   This sequence represents a fragment of the C4H enzyme coding sequence. It
CC   is an example of a DNA sequence of the invention, which are from
CC   Eucalyptus grandis (eucalyptus) and Pinus radiata (pine) associated with
CC   the lignin biosynthesis pathway. Constructs containing the DNA sequences
CC   can be used to produce transgenic plants or plant cells, especially woody
CC   plants e.g. eucalyptus or pine species but also e.g. monocotyledons or
CC   dicotyledons; by stably incorporating the constructs into the plant
CC   genome. The lignin content or structure, or activity of a specific enzyme

```
CC    in the plant, can therefore be modulated. Reductions in lignin content or
CC    changes in composition are useful in tree processing for paper. High
CC    lignin content results in energy- and chemical-intensive separation
CC    methods in order to obtain the pure cellulose fibre required. Reductions
CC    in lignin content may also be useful for forage crops, whilst increases
CC    or changes in composition may be desirable to increase the mechanical
CC    strength of wood, change its colour or increase its resistance to rot.
CC    The sequences are also useful as probes to isolate DNA sequences encoding
CC    enzymes involved in the lignin biosynthesis pathway from other plant
CC    species.
XX
SQ    Sequence 622 BP; 170 A; 117 C; 178 G; 157 T; 0 other;


   Query Match              13.8%;   Score 59.6;   DB 19;   Length 622;
   Best Local Similarity    54.1%;   Pred. No. 2.6e-08;
   Matches  119;  Conservative    0;  Mismatches  101;  Indels    0;  Gaps    0;


Qy     61 cctgctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcgcgctgcg 120
          ||  ||  | || || |||||||||| ||   | ||| || || |      || ||     |
Db    180 ccgactattgccgtttgggatggggaggagaagttgtcctggtgctggccttgccaatag 239

Qy    121 gaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacggttgatgg 180
             ||  | |||| || || ||||  |  | |||||||||| |||||     |||| |
Db    240 agtggtgagcttggtcctggcggcgcttattcagtgcttcgaatgggaacgagttggcga 299

Qy    181 agctcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcccgttgga 240
          ||     |||          |             || || || |||||   |   | || ||||||
Db    300 agaattggtggacttgtccgaggggacgggactcacaatgccaaagagagagccattgga 359

Qy    241 ggccatgtgcangccgcgtacagctatgcgtggtgttctt 280
          |||| ||||||     ||||| |    ||| | |  | ||
Db    360 ggccttgtgcaaagcgcgtgaatgcatgatagctaatgtt 399


RESULT    2
AAZ06838
ID    AAZ06838 standard; cDNA; 622 BP.
XX
AC    AAZ06838;
XX
DT    09-NOV-1999  (first entry)
XX
DE    Eucalyptus cinnamate 4-hydroxylase (C4H) cDNA.
XX
KW    Lignin; biosynthesis; forage crop; wood; paper production;
KW    transgenic plant; ss.
XX
OS    Eucalyptus grandis.
XX
PN    US5952486-A.
XX
PD    14-SEP-1999.
XX
PF    21-NOV-1997;  97US-0975316.
XX
PR    21-NOV-1997;  97US-0975316.
```

CC   This sequence represents a cinnamate 4-hydroxylase (C4H)
CC   cDNA from Eucalyptus grandis. This enzyme is involved in the
CC   biosynthesis of lignin, an insoluble polymer which is primarily
CC   responsible for the rigidity of plant stems. Lignin serves as a matrix
CC   around the polysaccharide components of some plant cell walls. The
CC   higher the lignin content, the more rigid the plant. Lignin also plays a
CC   role in disease resistance of plants by impeding the penetration and
CC   propagation of pathogenic agents. Lignin is formed by polymerisation of
CC   at least three different monolignols (para-coumaryl alcohol, coniferyl
CC   alcohol and sinapyl alcohol). These three monolignols are synthesised by
CC   similar pathways from phenylalanine in a multistep process and are
CC   believed to be polymerised into lignin via a free radical mechanism.
CC   The lignin content of plants can be altered using DNA sequences encoding
CC   these enzymes. Lignin content can be increased by incorporation of
CC   additional copies of genes encoding these enzymes into the target plant.
CC   This could be beneficial for increasing the mechanical strength of wood.
CC   Similarly, a decrease in lignin content can be obtained by transforming
CC   the target plant with antisense copies of such genes. This may be
CC   beneficial in plants used as forage crops for livestock (lignin is
CC   indigestible) and in trees used in paper manufacture.
XX
SQ   Sequence 622 BP; 170 A; 117 C; 178 G; 157 T; 0 other;

      Query Match   13.8%;   Score 59.6; DB 20;   Length 622;
      Best Local Similarity   54.1%;   Pred. No. 2.6e-08;
      Matches 119; Conservative   0; Mismatches 101; Indels   0; Gaps   0;

Qy   61 cctgctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcgcgctgcg 120
            || || | || || |||||||||| || | ||| || || | || || |
Db   180 ccgactattgccgtttgggatggggaggagaagttgtcctggtgctggccttgccaatag 239

Qy   121 gaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacggttgatgg 180
            || | ||||| || || |||| | | ||||||||| ||||| |||| |
Db   240 agtggtgagcttggtcctggcggcgcttattcagtgcttcgaatgggaacgagttggcga 299

Qy   181 agctcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcccgttgga 240
            || ·||| .| || || || ||||| | | || |||||·
Db   300 agaattggtggacttgtccgaggggacgggactcacaatgccaaagagagagccattgga 359

Qy   241 ggccatgtgcangccgcgtacagctatgcgtggtgttctt 280
            ||.|| |||||| ||||| | ||| | | | | ||
Db   360 ggccttgtgcaaagcgcgtgaatgcatgatagctaatgtt 399

```
RESULT   3
AAA67924
ID    AAA67924 standard; DNA; 622 BP.
XX
AC    AAA67924;
XX
DT    24-OCT-2000   (first entry)
XX
DE    Eucalyptus grandis C4H nucleotide sequence SEQ ID NO:17.
XX
KW    Plant; lignin; lignin biosynthetic pathway; Eucalyptus grandis;
KW    Pinus radiata; Monterey pine; ds.
XX
OS    Eucalyptus grandis.
XX
PN    WO200022099-A1.
XX
PD    20-APR-2000.
XX
PF    06-OCT-1999;   99WO-NZ00168.
XX
PR    09-OCT-1998;   98US-0169789.
PR    14-JUL-1999;   99US-0143811.
XX
PA    (GENE-) GENESIS RES & DEV CORP LTD.
PA    (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI    Bloksberg LN,  Havukkala IJ;
XX
DR    WPI; 2000-317962/27.
XX
PT    Novel polynucleotide encoding enzymes involved in lignin-biosynthetic
PT    pathway useful for producing transgenic plants especially eucalyptus
PT    and pine species having altered lignin content, composition and
PT    structure  -
XX
PS    Example 1; Page 61-62; 213pp; English.
XX
CC    The present invention describes isolated polynucleotides and proteins
CC    encoding and representing the enzymes cinnamate 4-hydroxylase (C4H),
CC    coumarate 3-hydroxylase (C3H), phenolase (PNL), O-methyl transferase
CC    (OMT), cinnamyl alcohol dehydrogenase (CAD), cinnamoyl-CoA reductase
CC    (CCR), phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL),
CC    coniferol glucosyl transferase (CGT), coniferin beta-glucosidase (CBG),
CC    laccase, peroxidase, ferulate-5-hydroxylase (F5H), alpha-amylase,
CC    caffeic acid methyl transferase, caffeoyl CoA methyl transferase,
CC    coumerate CoA ligase, cytochrome P450 LXX1A, diphenol oxidase, flavanol
CC    glucosyl transferase, flavenoid hydroxylase, and isoflavone reductase,
CC    which are involved in the lignin biosynthetic pathway. The
CC    polynucleotides can be used for modulating lignin content, lignin
CC    composition and the structure of a plant, especially eucalyptus and pine
CC    species, and for modifying the activity of an enzyme involved in lignin
CC    biosynthetic pathway, and for producing a plant having altered lignin
CC    content, composition and structure. They can be used for designing probes
CC    and primers useful for detecting similar DNA and RNA sequences in any
```

     Query Match   .   13.8%;   Score 59.6;   DB 21;   Length 622;
     Best Local Similarity   54.1%;   Pred. No. 2.6e-08;
     Matches 119; Conservative   0; Mismatches 101; Indels   0; Gaps   0;

Qy   61 cctgctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcgcgctgcg 120
           || || | || || ||||||||| || | ||| || || | || || |
Db   180 ccgactattgccgtttgggatggggaggagaagttgtcctggtgctggccttgccaatag 239

Qy   121 gaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacggttgatgg 180
           || | |||| || || |||| | | ||||||||| ||||| |||| |
Db   240 agtggtgagcttggtcctggcggcgcttattcagtgcttcgaatgggaacgagttggcga 299

Qy   181 agctcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcccgttgga 240
           || ||| | · || || || ||||| | | || |||||
Db   300 agaattggtggacttgtccgaggggacgggactcacaatgccaaagagagagccattgga 359

Qy   241 ggccatgtgcangccgcgtacagctatgcgtggtgttctt 280 ·
           ||||| |||||| ||||| | ||| | | | | ||
Db   360 ggccttgtgcaaagcgcgtgaatgcatgatagctaatgtt 399

RESULT   4
AAC47389
ID   AAC47389 standard; DNA; 1655 BP.
XX
AC   AAC47389;
XX
DT· 18-OCT-2000   (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 53631.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.

```
PR    01-APR-1999;    99US-0127462.
PR    06-APR-1999;    99US-0128234.
PR    08-APR-1999;    99US-0128714.
PR    16-APR-1999;    99US-0129845.
PR    19-APR-1999;    99US-0130077.
PR    21-APR-1999;    99US-0130449.
PR    23-APR-1999;    99US-0130510.
PR    23-APR-1999;    99US-0130891.
PR    28-APR-1999;    99US-0131449.
PR    30-APR-1999;    99US-0132048.
PR    30-APR-1999;    99US-0132407.
PR    04-MAY-1999;    99US-0132484.
PR    05-MAY-1999;    99US-0132485.
PR    06-MAY-1999;    99US-0132486.
PR    06-MAY-1999;    99US-0132487.
PR    07-MAY-1999;    99US-0132863.
PR    11-MAY-1999;    99US-0134256.
PR    14-MAY-1999;    99US-0134218.
PR    14-MAY-1999;    99US-0134219.
PR    14-MAY-1999;    99US-0134221.
PR    14-MAY-1999;    99US-0134370.
PR    18-MAY-1999;    99US-0134768.
PR    19-MAY-1999;    99US-0134941.
PR    20-MAY-1999;    99US-0135124.
PR    21-MAY-1999;    99US-0135353.
PR    24-MAY-1999;    99US-0135629.
PR    25-MAY-1999;    99US-0136021.
PR    27-MAY-1999;    99US-0136392.
PR    28-MAY-1999;    99US-0136782.
PR    01-JUN-1999;    99US-0137222.
PR    03-JUN-1999;    99US-0137528.
PR    04-JUN-1999;    99US-0137502.
PR    07-JUN-1999;    99US-0137724.
PR    08-JUN-1999;    99US-0138094.
PR    10-JUN-1999;    99US-0138540.
PR    10-JUN-1999;    99US-0138847.
PR    14-JUN-1999;    99US-0139119.
PR    16-JUN-1999;    99US-0139452.
PR    16-JUN-1999;    99US-0139453.
PR    17-JUN-1999;    99US-0139492.
PR    18-JUN-1999;    99US-0139454.
PR    18-JUN-1999;    99US-0139455.
PR    18-JUN-1999;    99US-0139456.
PR    18-JUN-1999;    99US-0139457.
PR    18-JUN-1999;    99US-0139458.
PR    18-JUN-1999;    99US-0139459.
PR    18-JUN-1999;    99US-0139460.
PR    18-JUN-1999;    99US-0139461.
PR    18-JUN-1999;    99US-0139462.
PR    18-JUN-1999;    99US-0139463.
PR    18-JUN-1999;    99US-0139750.
PR    18-JUN-1999;    99US-0139763.
PR    21-JUN-1999;    99US-0139817.
PR    22-JUN-1999;    99US-0139899.
PR    23-JUN-1999;    99US-0140353.
PR    23-JUN-1999;    99US-0140354.
PR    24-JUN-1999;    99US-0140695.
```

```
PR    28-JUN-1999;    99US-0140823.
PR    29-JUN-1999;    99US-0140991.
PR    30-JUN-1999;    99US-0141287.
PR    01-JUL-1999;    99US-0141842.
PR    01-JUL-1999;    99US-0142154.
PR    02-JUL-1999;    99US-0142055.
PR    06-JUL-1999;    99US-0142390.
PR    08-JUL-1999;    99US-0142803.
PR    09-JUL-1999;    99US-0142920.
PR    12-JUL-1999;    99US-0142977.
PR    13-JUL-1999;    99US-0143542.
PR    14-JUL-1999;    99US-0143624.
PR    15-JUL-1999;    99US-0144005.
PR    16-JUL-1999;    99US-0144085.
PR    16-JUL-1999;    99US-0144086.
PR    19-JUL-1999;    99US-0144325.
PR    19-JUL-1999;    99US-0144331.
PR    19-JUL-1999;    99US-0144332.
PR    19-JUL-1999;    99US-0144333.
PR    19-JUL-1999;    99US-0144334.
PR    19-JUL-1999;    99US-0144335.
PR    20-JUL-1999;    99US-0144352.
PR    20-JUL-1999;    99US-0144632.
PR    20-JUL-1999;    99US-0144884.
PR    21-JUL-1999;    99US-0144814.
PR    21-JUL-1999;    99US-0145086.
PR    21-JUL-1999;    99US-0145088.
PR    22-JUL-1999;    99US-0145085.
PR    22-JUL-1999;    99US-0145087.
PR    22-JUL-1999;    99US-0145089.
PR    22-JUL-1999;    99US-0145192.
PR    23-JUL-1999;    99US-0145145.
PR    23-JUL-1999;    99US-0145218.
PR    23-JUL-1999;    99US-0145224.
PR    26-JUL-1999;    99US-0145276.
PR    27-JUL-1999;    99US-0145913.
PR    27-JUL-1999;    99US-0145918.
PR    27-JUL-1999;    99US-0145919.
PR    28-JUL-1999;    99US-0145951.
PR    02-AUG-1999;    99US-0146386.
PR    02-AUG-1999;    99US-0146388.
PR    02-AUG-1999;    99US-0146389.
PR    03-AUG-1999;    99US-0147038.
PR    04-AUG-1999;    99US-0147204.
PR    04-AUG-1999;    99US-0147302.
PR    05-AUG-1999;    99US-0147192.
PR    05-AUG-1999;    99US-0147260.
PR    06-AUG-1999;    99US-0147303.
PR    06-AUG-1999;    99US-0147416.
PR    09-AUG-1999;    99US-0147493.
PR    09-AUG-1999;    99US-0147935.
PR    10-AUG-1999;    99US-0148171.
PR    11-AUG-1999;    99US-0148319.
PR    12-AUG-1999;    99US-0148341.
PR    13-AUG-1999;    99US-0148565.
PR    13-AUG-1999;    99US-0148684.
PR    16-AUG-1999;    99US-0149368.
```

```
PR    17-AUG-1999;    99US-0149175.
PR    18-AUG-1999;    99US-0149426.
PR    20-AUG-1999;    99US-0149722.
PR    20-AUG-1999;    99US-0149723.
PR    20-AUG-1999;    99US-0149929.
PR    23-AUG-1999;    99US-0149902.
PR    23-AUG-1999;    99US-0149930.
PR    25-AUG-1999;    99US-0150566.
PR    26-AUG-1999;    99US-0150884.
PR    27-AUG-1999;    99US-0151065.
PR    27-AUG-1999;    99US-0151066.
PR    27-AUG-1999;    99US-0151080.
PR    30-AUG-1999;    99US-0151303.
PR    31-AUG-1999;    99US-0151438.
PR    01-SEP-1999;    99US-0151930.
PR    07-SEP-1999;    99US-0152363.
PR    10-SEP-1999;    99US-0153070.
PR    13-SEP-1999;    99US-0153758.
PR    15-SEP-1999;    99US-0154018.
PR    16-SEP-1999;    99US-0154039.
PR    20-SEP-1999;    99US-0154779.
PR    22-SEP-1999;    99US-0155139.
PR    23-SEP-1999;    99US-0155486.
PR    24-SEP-1999;    99US-0155659.
PR    28-SEP-1999;    99US-0156458.
PR    29-SEP-1999;    99US-0156596.
PR    04-OCT-1999;    99US-0157117.
PR    05-OCT-1999;    99US-0157753.
PR    06-OCT-1999;    99US-0157865.
PR    07-OCT-1999;    99US-0158029.
PR    08-OCT-1999;    99US-0158232.
PR    12-OCT-1999;    99US-0158369.
PR    13-OCT-1999;    99US-0159293.
PR    13-OCT-1999;    99US-0159294.
PR    13-OCT-1999;    99US-0159295.
PR    14-OCT-1999;    99US-0159329.
PR    14-OCT-1999;    99US-0159330.
PR    14-OCT-1999;    99US-0159331.
PR    14-OCT-1999;    99US-0159637.
PR    14-OCT-1999;    99US-0159638.
PR    18-OCT-1999;    99US-0159584.
PR    21-OCT-1999;    99US-0160741.
PR    21-OCT-1999;    99US-0160767.
PR    21-OCT-1999;    99US-0160768.
PR    21-OCT-1999;    99US-0160770.
PR    21-OCT-1999;    99US-0160814.
PR    21-OCT-1999;    99US-0160815.
PR    22-OCT-1999;    99US-0160980.
PR    22-OCT-1999;    99US-0160981.
PR    22-OCT-1999;    99US-0160989.
PR    25-OCT-1999;    99US-0161404.
PR    25-OCT-1999;    99US-0161405.
PR    25-OCT-1999;    99US-0161406.
PR    26-OCT-1999;    99US-0161359.
PR    26-OCT-1999;    99US-0161360.
PR    26-OCT-1999;    99US-0161361.
PR    28-OCT-1999;    99US-0161920.
```

Query Match   13.7%;   Score 59;   DB 21;   Length 1655;
Best Local Similarity   52.7%;   Pred. No. 6e-08;
Matches 125; Conservative   0; Mismatches 112; Indels   0; Gaps   0;

```
Qy      32 gatggctccggcggcaaggccaaggggcccctgctgatccctttcgggatggggcggccc 91
           || || | ||| || |   | | |     |   |||||| || || |||| || |||
Db    1308 gacggatgcggaagcgattactatgtttacaagctgatgccgtttgggaatggccggaga 1367


Qy      92 aattgccccggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctc 151
           | ||| |||||| |     | | | | |||  ||     ||| ||| | | || || |
Db    1368 acttgtcccggcgccggattaggtcagaggattgtgacattggcgcttggatcgttgatt 1427


Qy     152 aattgcttcgactgggacacggttgatggagctcaggtttgacatgaagctancggcggg 211
           | |||||| || ||||| | || | || | |||   || |          | |    |||
Db    1428 caatgctttgaatgggagaatgtgaaaggggaagagatggatatgtctgagagtactggg 1487


Qy     212 ctgaccatgccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatg 268
           ||  ||||     |   |   || ||  |||||||||| | |||  |   |   ||||
Db    1488 ttgggtatgcgtaagatggatcctttacgggccatgtgtaggcctaggcccattatg 1544
```


RESULT   5
AAC37476

```
PR    19-APR-1999;    99US-0130077.
PR    21-APR-1999;    99US-0130449.
PR    23-APR-1999;    99US-0130510.
PR    23-APR-1999;    99US-0130891.
PR    28-APR-1999;    99US-0131449.
PR    30-APR-1999;    99US-0132048.
PR    30-APR-1999;    99US-0132407.
PR    04-MAY-1999;    99US-0132484.
PR    05-MAY-1999;    99US-0132485.
PR    06-MAY-1999;    99US-0132486.
PR    06-MAY-1999;    99US-0132487.
PR    07-MAY-1999;    99US-0132863.
PR    11-MAY-1999;    99US-0134256.
PR    14-MAY-1999;    99US-0134218.
PR    14-MAY-1999;    99US-0134219.
PR    14-MAY-1999;    99US-0134221.
PR    14-MAY-1999;    99US-0134370.
PR    18-MAY-1999;    99US-0134768.
PR    19-MAY-1999;    99US-0134941.
PR    20-MAY-1999;    99US-0135124.
PR    21-MAY-1999;    99US-0135353.
PR    24-MAY-1999;    99US-0135629.
PR    25-MAY-1999;    99US-0136021.
PR    27-MAY-1999;    99US-0136392.
PR    28-MAY-1999;    99US-0136782.
PR    01-JUN-1999;    99US-0137222.
PR    03-JUN-1999;    99US-0137528.
PR    04-JUN-1999;    99US-0137502.
PR    07-JUN-1999;    99US-0137724.
PR    08-JUN-1999;    99US-0138094.
PR    10-JUN-1999;    99US-0138540.
PR    10-JUN-1999;    99US-0138847.
PR    14-JUN-1999;    99US-0139119.
PR    16-JUN-1999;    99US-0139452.
PR    16-JUN-1999;    99US-0139453.
PR    17-JUN-1999;    99US-0139492.
PR    18-JUN-1999;    99US-0139454.
PR    18-JUN-1999;    99US-0139455.
PR    18-JUN-1999;    99US-0139456.
PR    18-JUN-1999;    99US-0139457.
PR    18-JUN-1999;    99US-0139458.
PR    18-JUN-1999;    99US-0139459.
PR    18-JUN-1999;    99US-0139460.
PR    18-JUN-1999;    99US-0139461.
PR    18-JUN-1999;    99US-0139462.
PR    18-JUN-1999;    99US-0139463.
PR    18-JUN-1999;    99US-0139750.
PR    18-JUN-1999;    99US-0139763.
PR    21-JUN-1999;    99US-0139817.
PR    22-JUN-1999;    99US-0139899.
PR    23-JUN-1999;    99US-0140353.
PR    23-JUN-1999;    99US-0140354.
PR    24-JUN-1999;    99US-0140695.
PR    28-JUN-1999;    99US-0140823.
PR    29-JUN-1999;    99US-0140991.
PR    30-JUN-1999;    99US-0141287.
PR    01-JUL-1999;    99US-0141842.
```

```
PR    01-JUL-1999;    99US-0142154.
PR    02-JUL-1999;    99US-0142055.
PR    06-JUL-1999;    99US-0142390.
PR    08-JUL-1999;    99US-0142803.
PR    09-JUL-1999;    99US-0142920.
PR    12-JUL-1999;    99US-0142977.
PR    13-JUL-1999;    99US-0143542.
PR    14-JUL-1999;    99US-0143624.
PR    15-JUL-1999;    99US-0144005.
PR    16-JUL-1999;    99US-0144085.
PR    16-JUL-1999;    99US-0144086.
PR    19-JUL-1999;    99US-0144325.
PR    19-JUL-1999;    99US-0144331.
PR    19-JUL-1999;    99US-0144332.
PR    19-JUL-1999;    99US-0144333.
PR    19-JUL-1999;    99US-0144334.
PR    19-JUL-1999;    99US-0144335.
PR    20-JUL-1999;    99US-0144352.
PR    20-JUL-1999;    99US-0144632.
PR    20-JUL-1999;    99US-0144884.
PR    21-JUL-1999;    99US-0144814.
PR    21-JUL-1999;    99US-0145086.
PR    21-JUL-1999;    99US-0145088.
PR    22-JUL-1999;    99US-0145085.
PR    22-JUL-1999;    99US-0145087.
PR    22-JUL-1999;    99US-0145089.
PR    22-JUL-1999;    99US-0145192.
PR    23-JUL-1999;    99US-0145145.
PR    23-JUL-1999;    99US-0145218.
PR    23-JUL-1999;    99US-0145224.
PR    26-JUL-1999;    99US-0145276.
PR    27-JUL-1999;    99US-0145913.
PR    27-JUL-1999;    99US-0145918.
PR    27-JUL-1999;    99US-0145919.
PR    28-JUL-1999;    99US-0145951.
PR    02-AUG-1999;    99US-0146386.
PR    02-AUG-1999;    99US-0146388.
PR    02-AUG-1999;    99US-0146389.
PR    03-AUG-1999;    99US-0147038.
PR    04-AUG-1999;    99US-0147204.
PR    04-AUG-1999;    99US-0147302.
PR    05-AUG-1999;    99US-0147192.
PR    05-AUG-1999;    99US-0147260.
PR    06-AUG-1999;    99US-0147303.
PR    06-AUG-1999;    99US-0147416.
PR    09-AUG-1999;    99US-0147493.
PR    09-AUG-1999;    99US-0147935.
PR    10-AUG-1999;    99US-0148171.
PR    11-AUG-1999;    99US-0148319.
PR    12-AUG-1999;    99US-0148341.
PR    13-AUG-1999;    99US-0148565.
PR    13-AUG-1999;    99US-0148684.
PR    16-AUG-1999;    99US-0149368.
PR    17-AUG-1999;    99US-0149175.
PR    18-AUG-1999;    99US-0149426.
PR    20-AUG-1999;    99US-0149722.
PR    20-AUG-1999;    99US-0149723.
```

```
PR    20-AUG-1999;    99US-0149929.
PR    23-AUG-1999;    99US-0149902.
PR    23-AUG-1999;    99US-0149930.
PR    25-AUG-1999;    99US-0150566.
PR    26-AUG-1999;    99US-0150884.
PR    27-AUG-1999;    99US-0151065..
PR    27-AUG-1999;    99US-0151066.
PR    27-AUG-1999;    99US-0151080.
PR    30-AUG-1999;    99US-0151303.
PR    31-AUG-1999;    99US-0151438.
PR    01-SEP-1999;    99US-0151930.
PR    07-SEP-1999;    99US-0152363.
PR    10-SEP-1999;    99US-0153070.
PR    13-SEP-1999;    99US-0153758.
PR    15-SEP-1999;    99US-0154018.
PR    16-SEP-1999;    99US-0154039.
PR    20-SEP-1999;    99US-0154779.
PR    22-SEP-1999;    99US-0155139.
PR    23-SEP-1999;    99US-0155486.
PR    24-SEP-1999;    99US-0155659.
PR    28-SEP-1999;    99US-0156458.
PR    29-SEP-1999;    99US-0156596.
PR    04-OCT-1999;    99US-0157117.
PR    05-OCT-1999;    99US-0157753.
PR    06-OCT-1999;    99US-0157865.
PR    07-OCT-1999;    99US-0158029.
PR    08-OCT-1999;    99US-0158232.
PR    12-OCT-1999;    99US-0158369.
PR    13-OCT-1999;    99US-0159293.
PR    13-OCT-1999;    99US-0159294.
PR    13-OCT-1999;    99US-0159295.
PR    14-OCT-1999;    99US-0159329.
PR    14-OCT-1999;    99US-0159330.
PR    14-OCT-1999;    99US-0159331.
PR    14-OCT-1999;    99US-0159637.
PR    14-OCT-1999;    99US-0159638.
PR    18-OCT-1999;    99US-0159584.
PR    21-OCT-1999;    99US-0160741.
PR    21-OCT-1999;    99US-0160767.
PR    21-OCT-1999;    99US-0160768.
PR    21-OCT-1999;    99US-0160770.
PR    21-OCT-1999;    99US-0160814.
PR    21-OCT-1999;    99US-0160815.
PR    22-OCT-1999;    99US-0160980.
PR    22-OCT-1999;    99US-0160981.
PR    22-OCT-1999;    99US-0160989.
PR    25-OCT-1999;    99US-0161404.
PR    25-OCT-1999;    99US-0161405.
PR    25-OCT-1999;    99US-0161406.
PR    26-OCT-1999;    99US-0161359.
PR    26-OCT-1999;    99US-0161360.
PR    26-OCT-1999;    99US-0161361.
PR    28-OCT-1999;    99US-0161920.
PR    28-OCT-1999;    99US-0161992.
PR    28-OCT-1999;    99US-0161993.
PR    29-OCT-1999;    99US-0162142.
```

```
Query Match              13.4%;   Score 58;   DB 21;   Length 1656;
Best Local Similarity    52.5%;   Pred. No. 1.2e-07;
Matches  124;  Conservative    0;  Mismatches 112;  Indels    0;  Gaps    0;
```

```
Qy     32 gatggctccggcggcaaggccaaggggcccctgctgatccctttcgggatggggcggccc 91
          || || | ||| || |   | | |    |  ||||||| || || |||| || |||
Db   1308 gacggatgcggaagcgattactatgtttacaagctgatgccgtttgggaatggccggaga 1367

Qy     92 aattgccccggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctc 151
          | ||| ||||||| |       | |  | | ||| ||      ||| ||| | | || || |
Db   1368 acttgtcccggcgccggattaggtcagaggattgtgacattggcgcttggatcgttgatt 1427

Qy    152 aattgcttcgactgggacacggttgatggagctcaggtttgacatgaagctancggcggg 211
          | ||||||| || |||||| |  || | || |   || |             | |   |||
Db   1428 caatgctttgaatgggagaatgtgaaaggggaagagatggatatgtctgagagtactggg 1487

Qy    212 ctgaccatgccccgggccgtcccgttggaggccatgtgcangccgcgtacagctat 267
          ||   ||||    |  | | || ||  |||||||||| | |||  | |    |||
Db   1488 ttgggtatgcgtaagatggatcctttacgggccatgtgtaggcctaggcccattat 1543
```

RESULT   6
AAX58406
ID   AAX58406 standard; cDNA; 1674 BP.
XX
AC   AAX58406;
XX
DT   02-AUG-1999 (first entry)
XX
DE   Jerusalem artichoke in-chain hydroxylase CYP81B1 clone D.
XX
KW   In-chain hydroxylase; transgenic plant; lipid; hydroxylation;
KW   oilseed; vegetable oil; crop protection; Jerusalem artichoke;
KW   CYP81B1; cytochrome P450; ss.
XX
OS   Helianthus tuberosus.
XX
FH   Key   Location/Qualifiers
FT   CDS   14..1531
FT   /*tag= a
XX
PN   WO9918224-A1.
XX
PD   15-APR-1999.
XX
PF   06-OCT-1998;   98WO-IB01716.
XX
PR   06-OCT-1997;   97US-0060960.
XX
PA   (CNRS) CENT NAT RECH SCI.
XX
PI   Batard Y, Benveniste I, Cabello-Huartado F, Durst F;
PI   Helvig C, Le Bouquin R, Pinot F, Salaun J, Tijet N;
PI   Werck-Reichhart D;
XX
DR   WPI; 1999-264030/22.
DR   P-PSDB; AAY05902.

```
XX
PT    Nucleic acid encoding plant fatty acid hydroxylases
XX
PS    Example 4; Fig 20A-B; 157pp; English.
XX
CC    This is the DNA sequence of clone D encoding in-chain hydroxylase
CC    CYP81B1 (see AAY05902) of Jerusalem artichoke.  Clone D was isolated
CC    from a tuber tissue cDNA library by PCR amplification.  CYP81B1
CC    is a microsomal cytochrome P450 that catalyses the omega-2, omega-3
CC    and omega-4 hydroxylation of capric, lauric and myristic acids.
CC    The major metabolite is the omega-3-hydroxylated compound.  The
CC    invention provides isolated nucleic acids (see AAX58400-06) encoding
CC    plant fatty acid hydroxylases (see AAY05896-902).  Also claimed are
CC    host cells, transgenic plants and compositions consisting of the
CC    plant fatty acid hydroxylase, a process for isolating additional
CC    fatty acid hydroxylase genes from a plant, and a process of
CC    altering fatty acid composition in a plant by expressing the plant
CC    fatty acid hydroxylase in a transgenic plant, and hydroxylating or
CC    epoxidating a fatty acid substrate in the plant.  Manipulating the
CC    hydroxylated fatty acid content of plants will modify resistance to
CC    drought and attack by insects and other pests.  The transgenic
CC    plants may also be used as sources of hydroxylated and epoxidized
CC    fatty acids useful in the manufacture of e.g. lubricants, anti-slip
CC    agents, plasticisers, coating agents, detergents and surfactants.
XX
SQ    Sequence 1674 BP; 427 A; 364 C; 404 G; 479 T; 0 other;


   Query Match             13.1%;   Score 56.6;  DB 20;   Length 1674;
   Best Local Similarity   49.8%;   Pred. No. 3.2e-07;
   Matches  140;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy      12 cggacttacccggccggttcgatggctccggcggcaaggccaagggcccctgctgatcc 71
           ||  |  ||| |  |||| || || ||`|  | ||| |     | ||| | | |  | |
Db    1251 cgttcaaaccagaaaggtttgaagggttagaagggacacgggatgggtttaagttattgc 1310

Qy      72 ctttcgggatggggcggcccaattgccccggggaaacgctcgcgctgcggaccgtcgggc 131
           | || |||   ||  ||   | ||| || ||||||    | ||| | || |   | ||| 
Db    1311 catttgggtctggaaggaggagttgtcctggggaaggcttggcggttcgaatgcttggga 1370

Qy     132 tggtgctcgcaacgctgctcaattgcttcgactgggacacggttgatggagctcaggttt 191
           ||    | |   | | | | | ||||||||| |||||        || ||    ||||
Db    1371 tgactttagggtcaattattcaatgcttcgattgggaacgaacgagtgaagagttggttg 1430

Qy     192 gacatgaagctancggcgggctgaccatgccccgggccgtcccgttggaggccatgtgca 251
                    |          ||||| |||||||||   ||| | || |||| || | |||||
Db    1431 atatgactgaaggtcctgggctaaccatgcctaaggctataccattggtagctaagtgca 1490

Qy     252 ngccgcgtacagctatgcgtggtgttcttaagaggctctga 292
             || ||    |  ||     | | ||        || |||
Db    1491 aacctcgggttgagatgacgaatctactgtccgaactgtga 1531


RESULT    7
AAA29326
ID    AAA29326 standard; cDNA; 1859 BP.
XX
```

```
AC    AAA29326;
XX
DT    26-SEP-2000   (first entry)
XX
DE    Soybean isoflavone-2-hydroxylase coding sequence.
XX
KW    Soybean; isoflavone-2-hydroxylase; flavonol; biosynthesis; anthocyanin;
KW    flower colour; pollen tube; feeding deterrent; UV irradiation; ss.
XX
OS    Glycine max.
XX
FH    Key             Location/Qualifiers
FT    CDS             59..1561
FT                    /*tag=  a
FT                    /product= Isoflavone-2-hydroxylase
XX
PN    WO200037652-A2.
XX
PD    29-JUN-2000.
XX
PF    20-DEC-1999;    99WO-US30337.
XX
PR    21-DEC-1998;    98US-0113190.
XX
PA    (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI    Famodu OO,  McGonigle B,  Odell JT,  Fader GM,  Falco SC;
XX
DR    WPI; 2000-442678/38.
DR    P-PSDB; AAY96593.
XX
PT    New polynucleotide encoding flavonoid biosynthetic enzymes, useful for
PT    producing transgenic plants and immunological screening of cDNA
PT    libraries
XX
PS    Claim 1; Page 30; 36pp; English.
XX
CC    This cDNA, isolated from clone sls1c.pk005.n3, encodes a plant (soybean)
CC    isoflavone-2-hydroxylase. It was determined using the sequence of an
CC    isoflavone-2-hydroxylase encoded by a contig composed of clones
CC    sgc1c.pk001.g17, sgs2c.pk004.h7 and slf1.pk0034.g1. The cDNA sequences
CC    can be used for the recombinant production of the enzyme, to isolate
CC    homologues, to create transgenic plants and to provide probes for
CC    genetically and physically mapping genes and as markers for traits linked
CC    to the genes. The proteins can be used for immunological screening, in
CC    particular to raise antibodies against the enzymes. The enzyme and its
CC    gene are useful to study flavonol biosynthesis in plants and provide
CC    means to enhance or otherwise alter flavonol and anthocyanin
CC    biosynthesis. Flavonoids have diverse functions, such as co-pigments in
CC    flower colour, stimulation of pollen tube growth, pollinator attraction,
CC    and feeding deterrents and protection against UV irradiation in fruits
CC    and seeds.
XX
SQ    Sequence 1859 BP; 536 A; 402 C; 417 G; 504 T; 0 other;
```

Query Match   12.4%; Score 53.4; DB 21; Length 1859;

```
      Best Local Similarity   53.1%;   Pred. No. 3.1e-06;
      Matches  111;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;


Qy      54 aggggcccctgctgatccctttcgggatggggcggcccaattgccccggggaaacgctcg 113
           ||| |      | || |  | || || ||||| |     |||||| || ||| |  | |
Db    1329 aggagaaaaagttggtagcatttggcatgggaagaagggcttgcccaggagaacccatgg 1388


Qy     114 cgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacgg 173
           |  ||| |  ||| |  |     | | | | || | || || |||||| |      |
Db    1389 ctatgcaaagtgtcagctttactttgggattgttgattcaatgttttgactggaaacgag 1448


Qy     174 ttgatggagctcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcc 233
           |   || |   || ||      || | |    || | ||| || |  ||     | |
Db    1449 taagtgaggaaaagcttgatatgacagagaacaattggatcaccttgtcaaggttaattc 1508


Qy     234 cgttggaggccatgtgcangccgcgtaca 262
           | |||||||||||||||||| | | ||  ||
Db    1509 cattggaggccatgtgcaaggctcgccca 1537



RESULT    8
AAQ50511
ID    AAQ50511 standard; cDNA; 1817 BP.
XX
AC    AAQ50511;
XX
DT    17-MAY-1994   (first entry)
XX
DE    Bxl gene.
XX
KW    Bxl; resistance; plant; benzoxazine; biosynthesis; allele;
KW    European corn borer; pest; vector; clone; ds.
XX
OS    Zea mays.
XX
FH    Key             Location/Qualifiers
FT    CDS             78..1670
FT                    /*tag=  a
FT                    /product= Bxl_gene_product
XX
PN    WO9322441-A.
XX
PD    11-NOV-1993.
XX
PF    23-APR-1992;   92WO-EP00905.
XX
PR    23-APR-1992;   92WO-EP00905.
XX
PA    (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
PI    Frey M,  Gierl A,  Peterson PA,  Saedler H,  Sommer H;
XX
DR    WPI; 1993-368800/46.
DR    P-PSDB; AAR43024.
XX
PT    DNA sequence of Bxl gene - used to confer resistance on plants
PT    with low or no levels of benzoxazine(s)
```

SQ   Sequence 1817 BP; 353 A; 607 C; 548 G; 309 T; 0 other;

   Query Match   11.9%;   Score 51.6;   DB 14;   Length 1817;
   Best Local Similarity   58.4%;   Pred. No. 1.1e-05;
   Matches   90; Conservative   0; Mismatches   64; Indels   0; Gaps   0;

Qy   19 acccggccggttcgatggctccggcggcaaggccaaggggcccctgctgatccctttcgg 78
           || |||||| || | || | | | ||||| || | | | || |||||
Db   1409 acaaggccgcgacgccgaggtcgacatgtacggcaaggacatccggttcgtgccgttcgg 1468

Qy   79 gatggggcggcccaattgccccggggaaacgctcgcgctgcggaccgtcgggctggtgct 138
           | ||||| | ||| | |||| ||| ||| | |||||||| | | ||||
Db   1469 ggctgggcgcaggatctgcgcgggggccacgttcgccatcgccaccgtcgagatcatgct 1528

Qy   139 cgcaacgctgctcaattgcttcgactgggacacg 172
           ||| | || || | |||||||||||| | |
Db   1529 cgcgaacctcatctaccatttcgactgggagatg 1562

RESULT   9
AAZ50024
ID   AAZ50024 standard; cDNA; 1847 BP.
XX
AC   AAZ50024;
XX
DT   25-APR-2000   (first entry)
XX
DE   Maize cytochrome p450 monooxygenase; CYP71C3v2 cDNA.
XX
KW   Cytochrome p450 monooxygenase; CYP71C3v2; maize; chromosome 4p; weed;
KW   p450 gene; molecular dioxygen; herbicidal; pigweed; transgenic organism;
KW   herbicide resistant; triasulfuron; quack grass; velvet leaf;
KW   labs quarter; Chenopodium album; ss.
XX
OS   Zea mays.
XX
FH   Key   Location/Qualifiers
FT   5'UTR   1..6
FT   /*tag=   a
FT   CDS   7..1611
FT   /*tag=   b
FT   /product= "Maize cytochrome p450 monooxygenase
FT   CYP71C3v2" misc_feature   540..541
FT   /*tag=   c
FT   /note= "intron 1 (AAZ50025) is located between these
FT   nucleotides"
FT   misc_feature   981..982
FT   /*tag=   d

```
FT                      /note= "intron 2 (AAZ50026) is located between these
FT                              nucleotides"
FT     polyA_signal     1678..1683
FT                      /*tag=  e
FT     polyA_signal     1700..1709
FT                      /*tag=  f
FT     polyA_signal     1728..1733
FT                      /*tag=  g
FT     polyA_signal     1763..1768
FT                      /*tag=  h
FT     polyA_signal     1806..1811
FT                      /*tag=  i
FT     polyA_site       1762
FT                      /*tag=  j
FT                      /note= "putative polyadenylation site"
FT     polyA_site       1799
FT                      /*tag=  k
FT                      /note= "putative polyadenylation site"
FT     polyA_site       1833
FT                      /*tag=  l
FT                      /note= "putative polyadenylation site"
XX
PN     WO200000502-A1.
XX
PD     06-JAN-2000.
XX
PF     23-JUN-1999;    99WO-US14117.
XX
PR     26-JUN-1998;    98US-0090759.
XX
PA     (UNII ) UNIV ILLINOIS FOUND.
XX
PI     Schuler MA,  Persans MW;
XX
DR     WPI; 2000-170902/15.
DR     P-PSDB; AAY44726.
XX
PT     Novel maize cytochrome P450 monooxygenase polypeptides and
PT     polynucleotides, used to confer triasulfuron herbicide resistance to
PT     plants  -
XX
PS     Claim 4; Page 46-48; 77pp; English.
XX
CC     The present sequence is the cDNA encoding maize cytochrome p450
CC     monooxygenase, CYP71C3v2. CYP71C3v2 gene is mapped to a single locus on
CC     the short arm of maize chromosome 4 (4p) and has two introns. It is
CC     encoded by a single copy or a small number of closely linked p450 genes.
CC     CYP71C3v2 reductively cleaves molecular dioxygen to produce
CC     functionalised organic substrates. It has herbicidal activity.
CC     CYP71C3v2 polynucleotides are used to produce transgenic organisms, such
CC     as yeast, plants and bacteria that are resistant to herbicides, such as
CC     triasulfurons. Undesired vegetation, e.g. weed, pigweed, velvet leaf,
CC     labs quarters, Chenopodium album and quack grass, can easily be
CC     controlled when such transgenic plants are grown. Transformed organisms
CC     can also be used to identify compounds with herbicidal activity.
XX
SQ     Sequence 1847 BP; 386 A; 576 C; 555 G; 330 T; 0 other;
```

```
    Query Match                11.4%;   Score 49.4;   DB 21;   Length 1847;
    Best Local Similarity      58.5%;   Pred. No. 5e-05;
    Matches    86;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy      22 cggccggttcgatggctccggcggcaaggccaaggggcccctgctgatccctttcgggat 81
            ||||  ||  |  |    |  |  ||  ||  |  |  |||      |  |      |  ||  ||||||
Db    1356 cggctgggacaagtccaacagctacagcggccaggacttcaggtacctgccgttcgggtc 1415

Qy      82 ggggcggcccaattgccccgggaaacgctcgcgctgcggaccgtcgggctggtgctcgc 141
            |||||  |   |   |  ||||||||||   |     |||||||     ||||  |  |  |  |  |||||||
Db    1416 tgggcgccggatctgccccggggccaacttcgcgctcgcgaccatggagatcatgctcgc 1475

Qy     142 aacgctgctcaattgcttcgactggga 168
            |   ||   |   |        |||||||||||
Db    1476 caacctcatgtaccatttcgactggga 1502


RESULT  10
AAZ87320
ID    AAZ87320 standard; cDNA; 1848 BP.
XX
AC    AAZ87320;
XX
DT    22-MAY-2000   (first entry)
XX
DE    Maize cytochrome P450 monooxygenase CYP71C3v2 full-length cDNA.
XX
KW    Cytochrome P450 monooxygenase; CYP71C3v2; herbicide detoxification;
KW    triasulfuron; transgenic plant; herbicide identification; ss.
XX
OS    Zea mays.
XX
FH    Key             Location/Qualifiers
FT    CDS             7..1611
FT                    /*tag=  a
FT                    /product= "Maize cytochrome P450 monooxygenase,
FT                              CYP71C3v2"
FT    exon            1..540
FT                    /*tag=  b
FT                    /number= 1
FT                    /note= "In genomic DNA, intron 1 (AAZ87321) lies between
FT                              exons 1 and 2"
FT    exon            541..981
FT                    /*tag=  c
FT                    /number= 2
FT                    /note= "In genomic DNA, intron 2 (AAZ87322) lies between
FT                              exons 2 and 3"
FT    exon            982..1847
FT                    /*tag=  c
FT                    /number= 3
XX
PN    WO200000585-A2.
XX
PD    06-JAN-2000.
XX
```

   Query Match   11.4%;   Score 49.4; DB 21;   Length 1848;
   Best Local Similarity   58.5%;   Pred. No. 5e-05;
   Matches   86; Conservative   0; Mismatches   61; Indels   0; Gaps   0;

Qy   22 cggccggttcgatggctccggcggcaaggccaagggggcccctgctgatccctttcgggat 81
           |||| || | | | | || || | | ||| | | | || ||||||
Db   1356 cggctgggacaagtccaacagctacagcggccaggacttcaggtacctgccgttcgggtc 1415

Qy   82 ggggcggcccaattgccccgggggaaacgctcgcgctgcggaccgtcgggctggtgctcgc 141
           ||||| | | |||||||||| | ||||||| |||| | | | | |||||||
Db   1416 tgggcgccggatctgccccggggccaacttcgcgctcgcgaccatggagatcatgctcgc 1475

Qy   142 aacgctgctcaattgcttcgactggga 168
           | || | | |||||||||||||
Db   1476 caacctcatgtaccatttcgactggga 1502

RESULT   11
AAC42545
ID   AAC42545 standard; DNA; 1545 BP.

```
XX
AC    AAC42545;
XX
DT    17-OCT-2000   (first entry)
XX
DE    Arabidopsis thaliana DNA fragment SEQ ID NO: 35961.
XX
KW    Hybridisation assay; genetic mapping; gene expression control;
KW    protein identification; signal transduction pathway;
KW    metabolic pathway; promoter; termination sequence; ss.
XX
OS    Arabidopsis thaliana.
XX
PN    EP1033405-A2.
XX
PD    06-SEP-2000.
XX
PF    25-FEB-2000; 2000EP-0301439.
XX
PR    25-FEB-1999;    99US-0121825.
PR    05-MAR-1999;    99US-0123180.
PR    09-MAR-1999;    99US-0123548.
PR    23-MAR-1999;    99US-0125788.
PR    25-MAR-1999;    99US-0126264.
PR    29-MAR-1999;    99US-0126785.
PR    01-APR-1999;    99US-0127462.
PR    06-APR-1999;    99US-0128234.
PR    08-APR-1999;    99US-0128714.
PR    16-APR-1999;    99US-0129845.
PR    19-APR-1999;    99US-0130077.
PR    21-APR-1999;    99US-0130449.
PR    23-APR-1999;    99US-0130510.
PR    23-APR-1999;    99US-0130891.
PR    28-APR-1999;    99US-0131449.
PR    30-APR-1999;    99US-0132048.
PR    30-APR-1999;    99US-0132407.
PR    04-MAY-1999;    99US-0132484.
PR    05-MAY-1999;    99US-0132485.
PR    06-MAY-1999;    99US-0132486.
PR    06-MAY-1999;    99US-0132487.
PR    07-MAY-1999;    99US-0132863.
PR    11-MAY-1999;    99US-0134256.
PR    14-MAY-1999;    99US-0134218.
PR    14-MAY-1999;    99US-0134219.
PR    14-MAY-1999;    99US-0134221.
PR    14-MAY-1999;    99US-0134370.
PR    18-MAY-1999;    99US-0134768.
PR    19-MAY-1999;    99US-0134941.
PR    20-MAY-1999;    99US-0135124.
PR    21-MAY-1999;    99US-0135353.
PR    24-MAY-1999;    99US-0135629.
PR    25-MAY-1999;    99US-0136021.
PR    27-MAY-1999;    99US-0136392.
PR    28-MAY-1999;    99US-0136782.
PR    01-JUN-1999;    99US-0137222.
PR    03-JUN-1999;    99US-0137528.
PR    04-JUN-1999;    99US-0137502.
```

```
PR    07-JUN-1999;    99US-0137724.
PR    08-JUN-1999;    99US-0138094.
PR    10-JUN-1999;    99US-0138540.
PR    10-JUN-1999;    99US-0138847.
PR    14-JUN-1999;    99US-0139119.
PR    16-JUN-1999;    99US-0139452.
PR    16-JUN-1999;    99US-0139453.
PR    17-JUN-1999;    99US-0139492.
PR    18-JUN-1999;    99US-0139454.
PR    18-JUN-1999;    99US-0139455.
PR    18-JUN-1999;    99US-0139456.
PR    18-JUN-1999;    99US-0139457.
PR    18-JUN-1999;    99US-0139458.
PR    18-JUN-1999;    99US-0139459.
PR    18-JUN-1999;    99US-0139460.
PR    18-JUN-1999;    99US-0139461.
PR    18-JUN-1999;    99US-0139462.
PR    18-JUN-1999;    99US-0139463.
PR    18-JUN-1999;    99US-0139750.
PR    18-JUN-1999;    99US-0139763.
PR    21-JUN-1999;    99US-0139817.
PR    22-JUN-1999;    99US-0139899.
PR    23-JUN-1999;    99US-0140353.
PR    23-JUN-1999;    99US-0140354.
PR    24-JUN-1999;    99US-0140695.
PR    28-JUN-1999;    99US-0140823.
PR    29-JUN-1999;    99US-0140991.
PR    30-JUN-1999;    99US-0141287.
PR    01-JUL-1999;    99US-0141842.
PR    01-JUL-1999;    99US-0142154.
PR    02-JUL-1999;    99US-0142055.
PR    06-JUL-1999;    99US-0142390.
PR    08-JUL-1999;    99US-0142803.
PR    09-JUL-1999;    99US-0142920.
PR    12-JUL-1999;    99US-0142977.
PR    13-JUL-1999;    99US-0143542.
PR    14-JUL-1999;    99US-0143624.
PR    15-JUL-1999;    99US-0144005.
PR    16-JUL-1999;    99US-0144085.
PR    16-JUL-1999;    99US-0144086.
PR    19-JUL-1999;    99US-0144325.
PR    19-JUL-1999;    99US-0144331.
PR    19-JUL-1999;    99US-0144332.
PR    19-JUL-1999;    99US-0144333.
PR    19-JUL-1999;    99US-0144334.
PR    19-JUL-1999;    99US-0144335.
PR    20-JUL-1999;    99US-0144352.
PR    20-JUL-1999;    99US-0144632.
PR    20-JUL-1999;    99US-0144884.
PR    21-JUL-1999;    99US-0144814.
PR    21-JUL-1999;    99US-0145086.
PR    21-JUL-1999;    99US-0145088.
PR    22-JUL-1999;    99US-0145085.
PR    22-JUL-1999;    99US-0145087.
PR    22-JUL-1999;    99US-0145089.
PR    22-JUL-1999;    99US-0145192.
PR    23-JUL-1999;    99US-0145145.
```

```
PR    23-JUL-1999;    99US-0145218.
PR    23-JUL-1999;    99US-0145224.
PR    26-JUL-1999;    99US-0145276.
PR    27-JUL-1999;    99US-0145913.
PR    27-JUL-1999;    99US-0145918.
PR    27-JUL-1999;    99US-0145919.
PR    28-JUL-1999;    99US-0145951.
PR    02-AUG-1999;    99US-0146386.
PR    02-AUG-1999;    99US-0146388.
PR    02-AUG-1999;    99US-0146389.
PR    03-AUG-1999;    99US-0147038.
PR    04-AUG-1999;    99US-0147204.
PR    04-AUG-1999;    99US-0147302.
PR    05-AUG-1999;    99US-0147192.
PR    05-AUG-1999;    99US-0147260.
PR    06-AUG-1999;    99US-0147303.
PR    06-AUG-1999;    99US-0147416.
PR    09-AUG-1999;    99US-0147493.
PR    09-AUG-1999;    99US-0147935.
PR    10-AUG-1999;    99US-0148171.
PR    11-AUG-1999;    99US-0148319.
PR    12-AUG-1999;    99US-0148341.
PR    13-AUG-1999;    99US-0148565.
PR    13-AUG-1999;    99US-0148684.
PR    16-AUG-1999;    99US-0149368.
PR    17-AUG-1999;    99US-0149175.
PR    18-AUG-1999;    99US-0149426.
PR    20-AUG-1999;    99US-0149722.
PR    20-AUG-1999;    99US-0149723.
PR    20-AUG-1999;    99US-0149929.
PR    23-AUG-1999;    99US-0149902.
PR    23-AUG-1999;    99US-0149930.
PR    25-AUG-1999;    99US-0150566.
PR    26-AUG-1999;    99US-0150884.
PR    27-AUG-1999;    99US-0151065.
PR    27-AUG-1999;    99US-0151066.
PR    27-AUG-1999;    99US-0151080.
PR    30-AUG-1999;    99US-0151303.
PR    31-AUG-1999;    99US-0151438.
PR    01-SEP-1999;    99US-0151930.
PR    07-SEP-1999;    99US-0152363.
PR    10-SEP-1999;    99US-0153070.
PR    13-SEP-1999;    99US-0153758.
PR    15-SEP-1999;    99US-0154018.
PR    16-SEP-1999;    99US-0154039.
PR    20-SEP-1999;    99US-0154779.
PR    22-SEP-1999;    99US-0155139.
PR    23-SEP-1999;    99US-0155486.
PR    24-SEP-1999;    99US-0155659.
PR    28-SEP-1999;    99US-0156458.
PR    29-SEP-1999;    99US-0156596.
PR    04-OCT-1999;    99US-0157117.
PR    05-OCT-1999;    99US-0157753.
PR    06-OCT-1999;    99US-0157865.
PR    07-OCT-1999;    99US-0158029.
PR    08-OCT-1999;    99US-0158232.
PR    12-OCT-1999;    99US-0158369.
```

```
PR    13-OCT-1999;    99US-0159293.
PR    13-OCT-1999;    99US-0159294.
PR    13-OCT-1999;    99US-0159295.
PR    14-OCT-1999;    99US-0159329.
PR    14-OCT-1999;    99US-0159330.
PR    14-OCT-1999;    99US-0159331.
PR    14-OCT-1999;    99US-0159637.
PR    14-OCT-1999;    99US-0159638.
PR    18-OCT-1999;    99US-0159584.
PR    21-OCT-1999;    99US-0160741.
PR    21-OCT-1999;    99US-0160767.
PR    21-OCT-1999;    99US-0160768.
PR    21-OCT-1999;    99US-0160770.
PR    21-OCT-1999;    99US-0160814.
PR    21-OCT-1999;    99US-0160815.
PR    22-OCT-1999;    99US-0160980.
PR    22-OCT-1999;    99US-0160981.
PR    22-OCT-1999;    99US-0160989.
PR    25-OCT-1999;    99US-0161404.
PR    25-OCT-1999;    99US-0161405.
PR    25-OCT-1999;    99US-0161406.
PR    26-OCT-1999;    99US-0161359.
PR    26-OCT-1999;    99US-0161360.
PR    26-OCT-1999;    99US-0161361.
PR    28-OCT-1999;    99US-0161920.
PR    28-OCT-1999;    99US-0161992.
PR    28-OCT-1999;    99US-0161993.
PR    29-OCT-1999;    99US-0162142.


   Query Match          11.3%;   Score 48.8;  DB 21;   Length 1545;
   Best Local Similarity   52.5%;   Pred. No. 7e-05;
   Matches  104;  Conservative    0;  Mismatches   94;  Indels   0;  Gaps   0;


Qy     64 gctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcgcgctgcggac 123
          ||| ||| | || ||||||||| ||       ||| || || |     ||| || | |||
Db   1326 gctaatctcatttgggatgggacgaagagcttgtcctggagccgggctagctcatcggct 1385


Qy    124 cgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacggttgatggagc 183
          |   | || || | || || | ||   || |  | || || || ||||| |  |||| || |
Db   1386 aataaaccaggctcttggaagtttggttcaatgttttgagtgggaaagagttggtgagga 1445


Qy    184 tcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcccgttggaggc 243
          |   ||       |   |    ||   ||  |||||  || | || ||     ||
Db   1446 ttttgtggacatgaccgaagacaaaggagccacattgcccaaagctataccattaagagc 1505


Qy    244 catgtgcangccgcgtac 261
          ||||||||   | ||| |
Db   1506 catgtgcaaagcacgttc 1523



RESULT  12
AAC47416
ID    AAC47416 standard; DNA; 1576 BP.
XX
AC    AAC47416;
XX
DT    18-OCT-2000  (first entry)
```

```
XX
DE     Arabidopsis thaliana DNA fragment SEQ ID NO: 53738.
XX
KW     Hybridisation assay; genetic mapping; gene expression control;
KW     protein identification; signal transduction pathway;
KW     metabolic pathway; promoter; termination sequence; ss.
XX
OS     Arabidopsis thaliana.
XX
PN     EP1033405-A2.
XX
PD     06-SEP-2000.
XX
PF     25-FEB-2000; 2000EP-0301439.
XX
PR     25-FEB-1999;    99US-0121825.
PR     05-MAR-1999;    99US-0123180.
PR     09-MAR-1999;    99US-0123548.
PR     23-MAR-1999;    99US-0125788.
PR     25-MAR-1999;    99US-0126264.
PR     29-MAR-1999;    99US-0126785.
PR     01-APR-1999;    99US-0127462.
PR     06-APR-1999;    99US-0128234.
PR     08-APR-1999;    99US-0128714.
PR     16-APR-1999;    99US-0129845.
PR     19-APR-1999;    99US-0130077.
PR     21-APR-1999;    99US-0130449.
PR     23-APR-1999;    99US-0130510.
PR     23-APR-1999;    99US-0130891.
PR     28-APR-1999;    99US-0131449.
PR     30-APR-1999;    99US-0132048.
PR     30-APR-1999;    99US-0132407.
PR     04-MAY-1999;    99US-0132484.
PR     05-MAY-1999;    99US-0132485.
PR     06-MAY-1999;    99US-0132486.
PR     06-MAY-1999;    99US-0132487.
PR     07-MAY-1999;    99US-0132863.
PR     11-MAY-1999;    99US-0134256.
PR     14-MAY-1999;    99US-0134218.
PR     14-MAY-1999;    99US-0134219.
PR     14-MAY-1999;    99US-0134221.
PR     14-MAY-1999;    99US-0134370.
PR     18-MAY-1999;    99US-0134768.
PR     19-MAY-1999;    99US-0134941.
PR     20-MAY-1999;    99US-0135124.
PR     21-MAY-1999;    99US-0135353.
PR     24-MAY-1999;    99US-0135629.
PR     25-MAY-1999;    99US-0136021.
PR     27-MAY-1999;    99US-0136392.
PR     28-MAY-1999;    99US-0136782.
PR     01-JUN-1999;    99US-0137222.
PR     03-JUN-1999;    99US-0137528.
PR     04-JUN-1999;    99US-0137502.
PR     07-JUN-1999;    99US-0137724.
PR     08-JUN-1999;    99US-0138094.
PR     10-JUN-1999;    99US-0138540.
PR     10-JUN-1999;    99US-0138847.
```

```
PR    14-JUN-1999;    99US-0139119.
PR    16-JUN-1999;    99US-0139452.
PR    16-JUN-1999;    99US-0139453.
PR    17-JUN-1999;    99US-0139492.
PR    18-JUN-1999;    99US-0139454.
PR    18-JUN-1999;    99US-0139455.
PR    18-JUN-1999;    99US-0139456.
PR    18-JUN-1999;    99US-0139457.
PR    18-JUN-1999;    99US-0139458.
PR    18-JUN-1999;    99US-0139459.
PR    18-JUN-1999;    99US-0139460.
PR    18-JUN-1999;    99US-0139461.
PR    18-JUN-1999;    99US-0139462.
PR    18-JUN-1999;    99US-0139463.
PR    18-JUN-1999;    99US-0139750.
PR    18-JUN-1999;    99US-0139763.
PR    21-JUN-1999;    99US-0139817.
PR    22-JUN-1999;    99US-0139899.
PR    23-JUN-1999;    99US-0140353.
PR    23-JUN-1999;    99US-0140354.
PR    24-JUN-1999;    99US-0140695.
PR    28-JUN-1999;    99US-0140823.
PR.   29-JUN-1999;    99US-0140991.
PR    30-JUN-1999;.   99US-0141287.
PR    01-JUL-1999;    99US-0141842.
PR    01-JUL-1999;    99US-0142154.
PR    02-JUL-1999;    99US-0142055.
PR    06-JUL-1999;    99US-0142390.
PR    08-JUL-1999;    99US-0142803.
PR    09-JUL-1999;    99US-0142920.
PR    12-JUL-1999;    99US-0142977.
PR    13-JUL-1999;    99US-0143542.
PR    14-JUL-1999;    99US-0143624.
PR    15-JUL-1999;    99US-0144005.
PR    16-JUL-1999;    99US-0144085.
PR    16-JUL-1999;    99US-0144086.
PR    19-JUL-1999;    99US-0144325.
PR    19-JUL-1999;    99US-0144331.
PR    19-JUL-1999;    99US-0144332.
PR    19-JUL-1999;    99US-0144333.
PR    19-JUL-1999;    99US-0144334.
PR    19-JUL-1999;    99US-0144335.
PR    20-JUL-1999;    99US-0144352.
PR    20-JUL-1999;    99US-0144632.
PR    20-JUL-1999;    99US-0144884.
PR    21-JUL-1999;    99US-0144814.
PR    21-JUL-1999;    99US-0145086.
PR    21-JUL-1999;    99US-0145088.
PR    22-JUL-1999;    99US-0145085.
PR    22-JUL-1999;    99US-0145087.
PR    22-JUL-1999;    99US-0145089.
PR    22-JUL-1999;    99US-0145192.
PR    23-JUL-1999;    99US-0145145.
PR    23-JUL-1999;    99US-0145218.
PR    23-JUL-1999;    99US-0145224.
PR    26-JUL-1999;    99US-0145276.
PR    27-JUL-1999;    99US-0145913.
```

```
PR    27-JUL-1999;    99US-0145918.
PR    27-JUL-1999;    99US-0145919.
PR    28-JUL-1999;    99US-0145951.
PR    02-AUG-1999;    99US-0146386.
PR    02-AUG-1999;    99US-0146388.
PR    02-AUG-1999;    99US-0146389.
PR    03-AUG-1999;    99US-0147038.
PR    04-AUG-1999;    99US-0147204.
PR    04-AUG-1999;    99US-0147302.
PR    05-AUG-1999;    99US-0147192.
PR    05-AUG-1999;    99US-0147260.
PR    06-AUG-1999;    99US-0147303.
PR    06-AUG-1999;    99US-0147416.
PR    09-AUG-1999;    99US-0147493.
PR    09-AUG-1999;    99US-0147935.
PR    10-AUG-1999;    99US-0148171.
PR    11-AUG-1999;    99US-0148319.
PR    12-AUG-1999;    99US-0148341.
PR    13-AUG-1999;    99US-0148565.
PR    13-AUG-1999;    99US-0148684.
PR    16-AUG-1999;    99US-0149368.
PR    17-AUG-1999;    99US-0149175.
PR    18-AUG-1999;    99US-0149426.
PR    20-AUG-1999;    99US-0149722.
PR    20-AUG-1999;    99US-0149723.
PR    20-AUG-1999;    99US-0149929.
PR    23-AUG-1999;    99US-0149902.
PR    23-AUG-1999;    99US-0149930.
PR    25-AUG-1999;    99US-0150566.
PR    26-AUG-1999;    99US-0150884.
PR    27-AUG-1999;    99US-0151065.
PR    27-AUG-1999;    99US-0151066.
PR    27-AUG-1999;    99US-0151080.
PR    30-AUG-1999;    99US-0151303.
PR    31-AUG-1999;    99US-0151438.
PR    01-SEP-1999;    99US-0151930.
PR    07-SEP-1999;    99US-0152363.
PR    10-SEP-1999;    99US-0153070.
PR    13-SEP-1999;    99US-0153758.
PR    15-SEP-1999;    99US-0154018.
PR    16-SEP-1999;    99US-0154039.
PR    20-SEP-1999;    99US-0154779.
PR    22-SEP-1999;    99US-0155139.
PR    23-SEP-1999;    99US-0155486.
PR    24-SEP-1999;    99US-0155659.
PR    28-SEP-1999;    99US-0156458.
PR    29-SEP-1999;    99US-0156596.
PR    04-OCT-1999;    99US-0157117.
PR    05-OCT-1999;    99US-0157753.
PR    06-OCT-1999;    99US-0157865.
PR    07-OCT-1999;    99US-0158029.
PR    08-OCT-1999;    99US-0158232.
PR    12-OCT-1999;    99US-0158369.
PR    13-OCT-1999;    99US-0159293.
PR    13-OCT-1999;    99US-0159294.
PR    13-OCT-1999;    99US-0159295.
PR    14-OCT-1999;    99US-0159329.
```

```
PR    14-OCT-1999;    99US-0159330.
PR    14-OCT-1999;    99US-0159331.
PR    14-OCT-1999;    99US-0159637.
PR    14-OCT-1999;    99US-0159638.
PR    18-OCT-1999;    99US-0159584.
PR    21-OCT-1999;    99US-0160741. .
PR    21-OCT-1999;    99US-0160767.
PR    21-OCT-1999;    99US-0160768.
PR    21-OCT-1999;    99US-0160770.
PR    21-OCT-1999;    99US-0160814.
PR    21-OCT-1999;    99US-0160815.
PR    22-OCT-1999;    99US-0160980.
PR    22-OCT-1999;    99US-0160981.
PR    22-OCT-1999;    99US-0160989.
PR    25-OCT-1999;    99US-0161404.
PR    25-OCT-1999;    99US-0161405.
PR    25-OCT-1999;    99US-0161406.
PR    26-OCT-1999;    99US-0161359.
PR    26-OCT-1999;    99US-0161360.
PR    26-OCT-1999;    99US-0161361.
PR    28-OCT-1999;    99US-0161920.
PR    28-OCT-1999;    99US-0161992.
PR    28-OCT-1999;    99US-0161993.
PR    29-OCT-1999;    99US-0162142.

   Query Match              11.3%;   Score 48.8;   DB 21;   Length 1576;
   Best Local Similarity    50.0%;   Pred. No. 7.1e-05;
   Matches  119;  Conservative   0;  Mismatches  119;  Indels   0;  Gaps   0;


Qy     31 cgatggctccggcggcaaggccaagggggcccctgctgatccctttcgggatggggcggcc 90
          || |||    || | || |   | |  |    ||||||| | || || |  || |||
Db   1263 cggtggagaaggagaaaaagatgatgttcgtatgctgatagcgtttggaagcggacggag 1322

Qy     91 caattgccccggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgct 150
          |  || ||||||| |     || |||| || ||    | | | | | | || |   |
Db   1323 aatatgtcccggtgttggactagcgcacaagattgtgacattagcgttaggatcgttaat 1382

Qy    151 caattgcttcgactgggacacggttgatggagctcaggtttgacatgaagctancggcgg 210
          | |||||| || ||| | | ||| | | |    || ||          |          ||
Db   1383 tcaatgctttgattggaaaaaggtgaacgaaaaagagattgatatgagtgagggtccggg 1442

Qy    211 gctgaccatgccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatg 268
          | || | ||||    |   || |||||    ||| |||| | | | || |    |||
Db   1443 gatggctatgcgtatgatggtgccgttacgagccttgtgtaagactcgacccataatg 1500


RESULT  13
AAC35968
ID   AAC35968 standard; DNA; 1578 BP.
XX
AC   AAC35968;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 12072.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
```

```
KW    protein identification; signal transduction pathway;
KW    metabolic pathway; promoter; termination sequence; ss.
XX
OS    Arabidopsis thaliana.
XX
PN    EP1033405-A2.
XX
PD    06-SEP-2000.
XX
PF    25-FEB-2000; 2000EP-0301439.
XX
PR    25-FEB-1999;    99US-0121825.
PR    05-MAR-1999;    99US-0123180.
PR    09-MAR-1999;    99US-0123548.
PR    23-MAR-1999;    99US-0125788.
PR    25-MAR-1999;    99US-0126264.
PR    29-MAR-1999;    99US-0126785.
PR    01-APR-1999;    99US-0127462.
PR    06-APR-1999;    99US-0128234.
PR    08-APR-1999;    99US-0128714.
PR    16-APR-1999;    99US-0129845.
PR    19-APR-1999;    99US-0130077.
PR    21-APR-1999;    99US-0130449.
PR    23-APR-1999;    99US-0130510.
PR    23-APR-1999;    99US-0130891.
PR    28-APR-1999;    99US-0131449.
PR    30-APR-1999;    99US-0132048.
PR    30-APR-1999;    99US-0132407.
PR    04-MAY-1999;    99US-0132484.
PR    05-MAY-1999;    99US-0132485.
PR    06-MAY-1999;    99US-0132486.
PR    06-MAY-1999;    99US-0132487.
PR    07-MAY-1999;    99US-0132863.
PR    11-MAY-1999;    99US-0134256.
PR    14-MAY-1999;    99US-0134218.
PR    14-MAY-1999;    99US-0134219.
PR    14-MAY-1999;    99US-0134221.
PR    14-MAY-1999;    99US-0134370.
PR    18-MAY-1999;    99US-0134768.
PR    19-MAY-1999;    99US-0134941.
PR    20-MAY-1999;    99US-0135124.
PR    21-MAY-1999;    99US-0135353.
PR    24-MAY-1999;    99US-0135629.
PR    25-MAY-1999;    99US-0136021.
PR    27-MAY-1999;    99US-0136392.
PR    28-MAY-1999;    99US-0136782.
PR    01-JUN-1999;    99US-0137222.
PR    03-JUN-1999;    99US-0137528.
PR    04-JUN-1999;    99US-0137502.
PR    07-JUN-1999;    99US-0137724.
PR    08-JUN-1999;    99US-0138094.
PR    10-JUN-1999;    99US-0138540.
PR    10-JUN-1999;    99US-0138847.
PR    14-JUN-1999;    99US-0139119.
PR    16-JUN-1999;    99US-0139452.
PR    16-JUN-1999;    99US-0139453.
PR    17-JUN-1999;    99US-0139492.
```

```
PR    18-JUN-1999;    99US-0139454.
PR    18-JUN-1999;    99US-0139455.
PR    18-JUN-1999;    99US-0139456.
PR    18-JUN-1999;    99US-0139457.
PR    18-JUN-1999;    99US-0139458.
PR    18-JUN-1999;    99US-0139459.
PR    18-JUN-1999;    99US-0139460.
PR    18-JUN-1999;    99US-0139461.
PR    18-JUN-1999;    99US-0139462.
PR    18-JUN-1999;    99US-0139463.
PR    18-JUN-1999;    99US-0139750.
PR    18-JUN-1999;    99US-0139763.
PR    21-JUN-1999;    99US-0139817.
PR    22-JUN-1999;    99US-0139899.
PR    23-JUN-1999;    99US-0140353.
PR    23-JUN-1999;    99US-0140354.
PR    24-JUN-1999;    99US-0140695.
PR    28-JUN-1999;    99US-0140823.
PR    29-JUN-1999;    99US-0140991.
PR    30-JUN-1999;    99US-0141287.
PR    01-JUL-1999;    99US-0141842.
PR    01-JUL-1999;    99US-0142154.
PR    02-JUL-1999;    99US-0142055.
PR    06-JUL-1999;    99US-0142390.
PR    08-JUL-1999;    99US-0142803.
PR    09-JUL-1999;    99US-0142920.
PR    12-JUL-1999;    99US-0142977.
PR    13-JUL-1999;    99US-0143542.
PR    14-JUL-1999;    99US-0143624.
PR    15-JUL-1999;    99US-0144005.
PR    16-JUL-1999;    99US-0144085.
PR    16-JUL-1999;    99US-0144086.
PR    19-JUL-1999;    99US-0144325.
PR    19-JUL-1999;    99US-0144331.
PR    19-JUL-1999;    99US-0144332.
PR    19-JUL-1999;    99US-0144333.
PR    19-JUL-1999;    99US-0144334.
PR    19-JUL-1999;    99US-0144335.
PR    20-JUL-1999;    99US-0144352.
PR    20-JUL-1999;    99US-0144632.
PR    20-JUL-1999;    99US-0144884.
PR    21-JUL-1999;    99US-0144814.
PR    21-JUL-1999;    99US-0145086.
PR    21-JUL-1999;    99US-0145088.
PR    22-JUL-1999;    99US-0145085.
PR    22-JUL-1999;    99US-0145087.
PR    22-JUL-1999;    99US-0145089.
PR    22-JUL-1999;    99US-0145192.
PR    23-JUL-1999;    99US-0145145.
PR    23-JUL-1999;    99US-0145218.
PR    23-JUL-1999;    99US-0145224.
PR    26-JUL-1999;    99US-0145276.
PR    27-JUL-1999;    99US-0145913.
PR    27-JUL-1999;    99US-0145918.
PR    27-JUL-1999;    99US-0145919.
PR    28-JUL-1999;    99US-0145951.
PR    02-AUG-1999;    99US-0146386.
```

```
PR    02-AUG-1999;    99US-0146388.
PR    02-AUG-1999;    99US-0146389.
PR    03-AUG-1999;    99US-0147038.
PR    04-AUG-1999;    99US-0147204.
PR    04-AUG-1999;    99US-0147302.
PR    05-AUG-1999;    99US-0147192.
PR    05-AUG-1999;    99US-0147260.
PR    06-AUG-1999;    99US-0147303.
PR    06-AUG-1999;    99US-0147416.
PR    09-AUG-1999;    99US-0147493.
PR    09-AUG-1999;    99US-0147935.
PR    10-AUG-1999;    99US-0148171.
PR    11-AUG-1999;    99US-0148319.
PR    12-AUG-1999;    99US-0148341.
PR    13-AUG-1999;    99US-0148565.
PR    13-AUG-1999;    99US-0148684.
PR    16-AUG-1999;    99US-0149368.
PR    17-AUG-1999;    99US-0149175.
PR    18-AUG-1999;    99US-0149426.
PR    20-AUG-1999;    99US-0149722.
PR    20-AUG-1999;    99US-0149723.
PR    20-AUG-1999;    99US-0149929.
PR    23-AUG-1999;    99US-0149902.
PR    23-AUG-1999;    99US-0149930.
PR    25-AUG-1999;    99US-0150566.
PR    26-AUG-1999;    99US-0150884.
PR    27-AUG-1999;    99US-0151065.
PR    27-AUG-1999;    99US-0151066.
PR    27-AUG-1999;    99US-0151080.
PR    30-AUG-1999;    99US-0151303.
PR    31-AUG-1999;    99US-0151438.
PR    01-SEP-1999;    99US-0151930.
PR    07-SEP-1999;    99US-0152363.
PR    10-SEP-1999;    99US-0153070.
PR    13-SEP-1999;    99US-0153758.
PR    15-SEP-1999;    99US-0154018.
PR    16-SEP-1999;    99US-0154039.
PR    20-SEP-1999;    99US-0154779.
PR    22-SEP-1999;    99US-0155139.
PR    23-SEP-1999;    99US-0155486.
PR    24-SEP-1999;    99US-0155659.
PR    28-SEP-1999;    99US-0156458.
PR    29-SEP-1999;    99US-0156596.
PR    04-OCT-1999;    99US-0157117.
PR    05-OCT-1999;    99US-0157753.
PR    06-OCT-1999;    99US-0157865.
PR    07-OCT-1999;    99US-0158029.
PR    08-OCT-1999;    99US-0158232.
PR    12-OCT-1999;    99US-0158369.
PR    13-OCT-1999;    99US-0159293.
PR    13-OCT-1999;    99US-0159294.
PR    13-OCT-1999;    99US-0159295.
PR    14-OCT-1999;    99US-0159329.
PR    14-OCT-1999;    99US-0159330.
PR    14-OCT-1999;    99US-0159331.
PR    14-OCT-1999;    99US-0159637.
PR    14-OCT-1999;    99US-0159638.
```

```
PR     18-OCT-1999;     99US-0159584.
PR     21-OCT-1999;     99US-0160741.
PR     21-OCT-1999;     99US-0160767.
PR     21-OCT-1999;     99US-0160768.
PR     21-OCT-1999;     99US-0160770.
PR     21-OCT-1999;     99US-0160814.
PR     21-OCT-1999;     99US-0160815.
PR     22-OCT-1999;     99US-0160980.
PR     22-OCT-1999;     99US-0160981.
PR     22-OCT-1999;     99US-0160989.
PR     25-OCT-1999;     99US-0161404.
PR     25-OCT-1999;     99US-0161405.
PR     25-OCT-1999;     99US-0161406.
PR     26-OCT-1999;     99US-0161359.
PR     26-OCT-1999;     99US-0161360.
PR     26-OCT-1999;     99US-0161361.
PR     28-OCT-1999;     99US-0161920.
PR     28-OCT-1999;     99US-0161992.
PR     28-OCT-1999;     99US-0161993.
PR     29-OCT-1999;     99US-0162142.
```

```
   Query Match              11.3%;   Score 48.8;   DB 21;   Length 1578;
   Best Local Similarity   50.0%;   Pred. No. 7.1e-05;
   Matches  119;   Conservative    0;   Mismatches  119;   Indels    0;   Gaps    0;


Qy      31 cgatggctccggcggcaaggccaagggggcccctgctgatccctttcgggatggggcggcc 90
           ||  |||     ||  |   || |     | |   |   ||||||| | || || |  || |||
Db    1263 cggtggagaaggagaaaaagatgatgttcgtatgctgatagcgtttggaagcggacggag 1322


Qy      91 caattgccccgggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgct 150
            |  || ||||| |    || ||||     ||  ||     | | | | | | | ||| |  |
Db    1323 aatatgtcccggtgttggactagcgcacaagattgtgacattagcgttaggatcgttaat 1382


Qy     151 caattgcttcgactgggacacggttgatggagctcaggtttgacatgaagctancggcgg 210
             | ||||| || ||| | | ||| | | |     || ||  ·       |          ||
Db    1383 tcaatgctttgattggaaaaaggtgaacgaaaaagagattgatatgagtgagggtccggg 1442


Qy     211 gctgaccatgccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatg 268
           | || | |||     |    || ||||||   ||| |||| | | | || |    |||
Db    1443 gatggctatgcgtatgatggtgccgttacgagccttgtgtaagactcgacccataatg 1500
```


RESULT 14
AAC50087
ID   AAC50087 standard; DNA; 1497 BP.
XX
AC   AAC50087;
XX
DT   18-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 63546.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.

```
XX
PN    EP1033405-A2.
XX
PD    06-SEP-2000.
XX
PF    25-FEB-2000; 2000EP-0301439.
XX
PR    25-FEB-1999;    99US-0121825.
PR    05-MAR-1999;    99US-0123180.
PR    09-MAR-1999;    99US-0123548.
PR    23-MAR-1999;    99US-0125788.
PR    25-MAR-1999;    99US-0126264.
PR    29-MAR-1999;    99US-0126785.
PR    01-APR-1999;    99US-0127462.
PR    06-APR-1999;    99US-0128234.
PR    08-APR-1999;    99US-0128714.
PR    16-APR-1999;    99US-0129845.
PR    19-APR-1999;    99US-0130077.
PR    21-APR-1999;    99US-0130449.
PR    23-APR-1999;    99US-0130510.
PR    23-APR-1999;    99US-0130891.
PR    28-APR-1999;    99US-0131449.
PR    30-APR-1999;    99US-0132048.
PR    30-APR-1999;    99US-0132407.
PR    04-MAY-1999;    99US-0132484.
PR    05-MAY-1999;    99US-0132485.
PR    06-MAY-1999;    99US-0132486.
PR    06-MAY-1999;    99US-0132487.
PR    07-MAY-1999;    99US-0132863.
PR    11-MAY-1999;    99US-0134256.
PR    14-MAY-1999;    99US-0134218.
PR    14-MAY-1999;    99US-0134219.
PR    14-MAY-1999;    99US-0134221.
PR    14-MAY-1999;    99US-0134370.
PR    18-MAY-1999;    99US-0134768.
PR    19-MAY-1999;    99US-0134941.
PR    20-MAY-1999;    99US-0135124.
PR    21-MAY-1999;    99US-0135353.
PR    24-MAY-1999;    99US-0135629.
PR    25-MAY-1999;    99US-0136021.
PR    27-MAY-1999;    99US-0136392.
PR    28-MAY-1999;    99US-0136782.
PR    01-JUN-1999;    99US-0137222.
PR    03-JUN-1999;    99US-0137528.
PR    04-JUN-1999;    99US-0137502.
PR    07-JUN-1999;    99US-0137724.
PR    08-JUN-1999;    99US-0138094.
PR    10-JUN-1999;    99US-0138540.
PR    10-JUN-1999;    99US-0138847.
PR    14-JUN-1999;    99US-0139119.
PR    16-JUN-1999;    99US-0139452.
PR    16-JUN-1999;    99US-0139453.
PR    17-JUN-1999;    99US-0139492.
PR    18-JUN-1999;    99US-0139454.
PR    18-JUN-1999;    99US-0139455.
PR    18-JUN-1999;    99US-0139456.
PR    18-JUN-1999;    99US-0139457.
```

```
PR      18-JUN-1999;    99US-0139458.
PR      18-JUN-1999;    99US-0139459.
PR      18-JUN-1999;    99US-0139460.
PR      18-JUN-1999;    99US-0139461.
PR      18-JUN-1999;    99US-0139462.
PR      18-JUN-1999;    99US-0139463.
PR      18-JUN-1999;    99US-0139750.
PR      18-JUN-1999;    99US-0139763.
PR      21-JUN-1999;    99US-0139817.
PR      22-JUN-1999;    99US-0139899.
PR      23-JUN-1999;    99US-0140353.
PR      23-JUN-1999;    99US-0140354.
PR      24-JUN-1999;    99US-0140695.
PR      28-JUN-1999;    99US-0140823.
PR      29-JUN-1999;    99US-0140991.
PR      30-JUN-1999;    99US-0141287.
PR      01-JUL-1999;    99US-0141842.
PR      01-JUL-1999;    99US-0142154.
PR      02-JUL-1999;    99US-0142055.
PR      06-JUL-1999;    99US-0142390.
PR      08-JUL-1999;    99US-0142803.
PR      09-JUL-1999;    99US-0142920.
PR      12-JUL-1999;    99US-0142977.
PR      13-JUL-1999;    99US-0143542.
PR·     14-JUL-1999;    99US-0143624.
PR      15-JUL-1999;    99US-0144005.
PR      16-JUL-1999;    99US-0144085.
PR      16-JUL-1999;    99US-0144086.
PR      19-JUL-1999;    99US-0144325.
PR      19-JUL-1999;    99US-0144331.
PR      19-JUL-1999;    99US-0144332.
PR      19-JUL-1999;    99US-0144333.
PR      19-JUL-1999;    99US-0144334.
PR      19-JUL-1999;    99US-0144335.
PR      20-JUL-1999;    99US-0144352.
PR      20-JUL-1999;    99US-0144632.
PR      20-JUL-1999;    99US-0144884.
PR      21-JUL-1999;    99US-0144814.
PR      21-JUL-1999;    99US-0145086.
PR      21-JUL-1999;    99US-0145088.
PR      22-JUL-1999;    99US-0145085.
PR      22-JUL-1999;    99US-0145087.
PR      22-JUL-1999;    99US-0145089.
PR      22-JUL-1999;    99US-0145192.
PR      23-JUL-1999;    99US-0145145.
PR      ·23-JUL-1999;    99US-0145218.
PR      ·23-JUL-1999;    99US-0145224.
PR      26-JUL-1999;    99US-0145276.
PR      27-JUL-1999;    99US-0145913.
PR      27-JUL-1999;    99US-0145918.
PR      27-JUL-1999;    99US-0145919.
PR      28-JUL-1999;    99US-0145951.
PR      02-AUG-1999;    99US-0146386.
·PR     02-AUG-1999;    99US-0146388.
PR      02-AUG-1999;    99US-0146389.
PR      03-AUG-1999;    99US-0147038.
PR      04-AUG-1999;    99US-0147204.
```

```
PR    04-AUG-1999;    99US-0147302.
PR    05-AUG-1999;    99US-0147192.
PR    05-AUG-1999;    99US-0147260.
PR    06-AUG-1999;    99US-0147303.
PR    06-AUG-1999;    99US-0147416.
PR    09-AUG-1999;    99US-0147493.
PR    09-AUG-1999;    99US-0147935.
PR    10-AUG-1999;    99US-0148171.
PR    11-AUG-1999;    99US-0148319.
PR    12-AUG-1999;    99US-0148341.
PR    13-AUG-1999;    99US-0148565.
PR    13-AUG-1999;    99US-0148684.
PR    16-AUG-1999;    99US-0149368.
PR    17-AUG-1999;    99US-0149175.
PR    18-AUG-1999;    99US-0149426.
PR    20-AUG-1999;    99US-0149722.
PR    20-AUG-1999;    99US-0149723.
PR    20-AUG-1999;    99US-0149929.
PR    23-AUG-1999;    99US-0149902.
PR    23-AUG-1999;    99US-0149930.
PR    25-AUG-1999;    99US-0150566.
PR    26-AUG-1999;    99US-0150884.
PR    27-AUG-1999;    99US-0151065.
PR    .27-AUG-1999;   99US-0151066.
PR    27-AUG-1999;    99US-0151080.
PR    30-AUG-1999;    99US-0151303.
PR    31-AUG-1999;    99US-0151438.
PR    01-SEP-1999;    99US-0151930.
PR    07-SEP-1999;    99US-0152363.
PR    10-SEP-1999;    99US-0153070.
PR    13-SEP-1999;    99US-0153758.
PR    15-SEP-1999;    99US-0154018.
PR    16-SEP-1999;    99US-0154039.
PR    20-SEP-1999;    99US-0154779.
PR    22-SEP-1999;    99US-0155139.
PR    23-SEP-1999;    99US-0155486.
PR    24-SEP-1999;    99US-0155659.
PR    28-SEP-1999;    99US-0156458.
PR    29-SEP-1999;    99US-0156596.
PR    04-OCT-1999;    99US-0157117.
PR    05-OCT-1999;    99US-0157753.
PR    06-OCT-1999;    99US-0157865.
PR    07-OCT-1999;    99US-0158029.
PR    08-OCT-1999;    99US-0158232.
PR    12-OCT-1999;    99US-0158369.
PR    13-OCT-1999;    99US-0159293.
PR    13-OCT-1999;    99US-0159294.
PR    13-OCT-1999;    99US-0159295.
PR    14-OCT-1999;    99US-0159329.
PR    14-OCT-1999;    99US-0159330.
PR    14-OCT-1999;    99US-0159331.
PR    14-OCT-1999;    99US-0159637.
PR    14-OCT-1999;    99US-0159638.
PR    18-OCT-1999;    99US-0159584.
PR    21-OCT-1999;    99US-0160741.
PR    21-OCT-1999;    99US-0160767.
PR    21-OCT-1999;    99US-0160768.
```

    Query Match   10.8%;   Score 46.8; DB 21;   Length 1497;
    Best Local Similarity   52.0%;   Pred. No. 0.00028;
    Matches 102; Conservative   0; Mismatches   94; Indels   0; Gaps   0;

Qy   64 gctgatccctttcgggatggggcggcccaattgccccggggaaacgctcgcgctgcggac 123
           ||| || ||| |||||||| | || | || || || | || | ||||
Db   1266 gctaatgccgttcgggctaggaagaagggcatgtcctggatccggtttggctcagcggct 1325

Qy   124 cgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacggttgatggagc 183
           || || |||| | || | | || || || ||||| | | ||| | ||
Db   1326 agtgaccttgagtctcgggtctttgattcaatgtttttgaatgggagaggattggagaaga 1385

Qy   184 tcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcccgttggaggc 243
           |||| | || || || |||||| || |||||||| ||
Db   1386 agaggtcgatatgactgaaggtccaggactcacaatgccgaaagctagaccgttggaagc 1445

Qy   244 catgtgcangccgcgt 259
           |||||||| | |||
Db   1446 tatgtgcagagcacgt 1461

RESULT 15
AAC47053
ID   AAC47053 standard; DNA; 1519 BP.
XX
AC   AAC47053;
XX
DT   18-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 52389.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.

```
XX
PF    25-FEB-2000;  2000EP-0301439.
XX
PR    25-FEB-1999;   99US-0121825.
PR    05-MAR-1999;   99US-0123180.
PR    09-MAR-1999;   99US-0123548.
PR    23-MAR-1999;   99US-0125788.
PR    25-MAR-1999;   99US-0126264.
PR    29-MAR-1999;   99US-0126785.
PR    01-APR-1999;   99US-0127462.
PR    06-APR-1999;   99US-0128234.
PR    08-APR-1999;   99US-0128714.
PR    16-APR-1999;   99US-0129845.
PR    19-APR-1999;   99US-0130077.
PR    21-APR-1999;   99US-0130449.
PR    23-APR-1999;   99US-0130510.
PR    23-APR-1999;   99US-0130891.
PR    28-APR-1999;   99US-0131449.
PR    30-APR-1999;   99US-0132048.
PR    30-APR-1999;   99US-0132407.
PR    04-MAY-1999;   99US-0132484.
PR    05-MAY-1999;   99US-0132485.
PR    06-MAY-1999;   99US-0132486.
PR    06-MAY-1999;   99US-0132487.
PR    07-MAY-1999;   99US-0132863.
PR    11-MAY-1999;   99US-0134256.
PR    14-MAY-1999;   99US-0134218.
PR    14-MAY-1999;   99US-0134219.
PR    14-MAY-1999;   99US-0134221.
PR    14-MAY-1999;   99US-0134370.
PR    18-MAY-1999;   99US-0134768.
PR    19-MAY-1999;   99US-0134941.
PR    20-MAY-1999;   99US-0135124.
PR    21-MAY-1999;   99US-0135353.
PR    24-MAY-1999;   99US-0135629.
PR    25-MAY-1999;   99US-0136021.
PR    27-MAY-1999;   99US-0136392.
PR    28-MAY-1999;   99US-0136782.
PR    01-JUN-1999;   99US-0137222.
PR    03-JUN-1999;   99US-0137528.
PR    04-JUN-1999;   99US-0137502.
PR    07-JUN-1999;   99US-0137724.
PR    08-JUN-1999;   99US-0138094.
PR    10-JUN-1999;   99US-0138540.
PR    10-JUN-1999;   99US-0138847.
PR    14-JUN-1999;   99US-0139119.
PR    16-JUN-1999;   99US-0139452.
PR    16-JUN-1999;   99US-0139453.
PR    17-JUN-1999;   99US-0139492.
PR    18-JUN-1999;   99US-0139454.
PR    18-JUN-1999;   99US-0139455.
PR    18-JUN-1999;   99US-0139456.
PR    18-JUN-1999;   99US-0139457.
PR    18-JUN-1999;   99US-0139458.
PR    18-JUN-1999;   99US-0139459.
PR    18-JUN-1999;   99US-0139460.
PR    18-JUN-1999;   99US-0139461.
```

```
PR    18-JUN-1999;   99US-0139462.
PR    18-JUN-1999;   99US-0139463.
PR    18-JUN-1999;   99US-0139750.
PR    18-JUN-1999;   99US-0139763.
PR    21-JUN-1999;   99US-0139817.
PR    22-JUN-1999;   99US-0139899.
PR    23-JUN-1999;   99US-0140353.
PR    23-JUN-1999;   99US-0140354.
PR    24-JUN-1999;   99US-0140695.
PR    28-JUN-1999;   99US-0140823.
PR    29-JUN-1999;   99US-0140991.
PR    30-JUN-1999;   99US-0141287.
PR    01-JUL-1999;   99US-0141842.
PR    01-JUL-1999;   99US-0142154.
PR    02-JUL-1999;   99US-0142055.
PR    06-JUL-1999;   99US-0142390.
PR    08-JUL-1999;   99US-0142803.
PR    09-JUL-1999;   99US-0142920.
PR    12-JUL-1999;   99US-0142977.
PR    13-JUL-1999;   99US-0143542.
PR    14-JUL-1999;   99US-0143624.
PR    15-JUL-1999;   99US-0144005.
PR    16-JUL-1999;   99US-0144085.
PR    16-JUL-1999;   99US-0144086.
PR    19-JUL-1999;   99US-0144325.
PR    19-JUL-1999;   99US-0144331.
PR    19-JUL-1999;   99US-0144332.
PR    19-JUL-1999;   99US-0144333.
PR    19-JUL-1999;   99US-0144334.
PR    19-JUL-1999;   99US-0144335.
PR    20-JUL-1999;   99US-0144352.
PR    20-JUL-1999;   99US-0144632.
PR    20-JUL-1999;   99US-0144884.
PR    21-JUL-1999;   99US-0144814.
PR    21-JUL-1999;   99US-0145086.
PR    21-JUL-1999;   99US-0145088.
PR    22-JUL-1999;   99US-0145085.
PR    22-JUL-1999;   99US-0145087.
PR    22-JUL-1999;   99US-0145089.
PR    22-JUL-1999;   99US-0145192.
PR    23-JUL-1999;   99US-0145145.
PR    23-JUL-1999;   99US-0145218.
PR    23-JUL-1999;   99US-0145224.
PR    26-JUL-1999;   99US-0145276.
PR    27-JUL-1999;   99US-0145913.
PR    27-JUL-1999;   99US-0145918.
PR    27-JUL-1999;   99US-0145919.
PR    28-JUL-1999;   99US-0145951.
PR    02-AUG-1999;   99US-0146386.
PR    02-AUG-1999;   99US-0146388.
PR    02-AUG-1999;   99US-0146389.
PR    03-AUG-1999;   99US-0147038.
PR    04-AUG-1999;   99US-0147204.
PR    04-AUG-1999;   99US-0147302.
PR    05-AUG-1999;   99US-0147192.
PR    05-AUG-1999;   99US-0147260.
PR    06-AUG-1999;   99US-0147303.
```

```
PR    06-AUG-1999;    99US-0147416.
PR    09-AUG-1999;    99US-0147493.
PR    09-AUG-1999;    99US-0147935.
PR    10-AUG-1999;    99US-0148171.
PR    11-AUG-1999;    99US-0148319.
PR    12-AUG-1999;    99US-0148341.
PR    13-AUG-1999;    99US-0148565.
PR    13-AUG-1999;    99US-0148684.
PR    16-AUG-1999;    99US-0149368.
PR    17-AUG-1999;    99US-0149175.
PR    18-AUG-1999;    99US-0149426.
PR    20-AUG-1999;    99US-0149722.
PR    20-AUG-1999;    99US-0149723.
PR    20-AUG-1999;    99US-0149929.
PR    23-AUG-1999;    99US-0149902.
PR    23-AUG-1999;    99US-0149930.
PR    25-AUG-1999;    99US-0150566.
PR    26-AUG-1999;    99US-0150884.
PR    27-AUG-1999;    99US-0151065.
PR    27-AUG-1999;    99US-0151066.
PR    27-AUG-1999;    99US-0151080.
PR    30-AUG-1999;    99US-0151303.
PR    31-AUG-1999;    99US-0151438.
PR    01-SEP-1999;    99US-0151930.
PR    07-SEP-1999;    99US-0152363.
PR    10-SEP-1999;    99US-0153070.
PR    13-SEP-1999;    99US-0153758.
PR    15-SEP-1999;    99US-0154018.
PR    16-SEP-1999;    99US-0154039.
PR    20-SEP-1999;    99US-0154779.
PR    22-SEP-1999;    99US-0155139.
PR    23-SEP-1999;    99US-0155486.
PR    24-SEP-1999;    99US-0155659.
PR    28-SEP-1999;    99US-0156458.
PR    29-SEP-1999;    99US-0156596.
PR    04-OCT-1999;    99US-0157117.
PR    05-OCT-1999;    99US-0157753.
PR    06-OCT-1999;    99US-0157865.
PR    07-OCT-1999;    99US-0158029.
PR    08-OCT-1999;    99US-0158232.
PR    12-OCT-1999;    99US-0158369.
PR    13-OCT-1999;    99US-0159293.
PR    13-OCT-1999;    99US-0159294.
PR    13-OCT-1999;    99US-0159295.
PR    14-OCT-1999;    99US-0159329.
PR    14-OCT-1999;    99US-0159330.
PR    14-OCT-1999;    99US-0159331.
PR    14-OCT-1999;    99US-0159637.
PR    14-OCT-1999;    99US-0159638.
PR    18-OCT-1999;    99US-0159584.
PR    21-OCT-1999;    99US-0160741.
PR    21-OCT-1999;    99US-0160767.
PR    21-OCT-1999;    99US-0160768.
PR    21-OCT-1999;    99US-0160770.
PR    21-OCT-1999;    99US-0160814.
PR    21-OCT-1999;    99US-0160815.
PR    22-OCT-1999;    99US-0160980.
```

```
PR    22-OCT-1999;    99US-0160981.
PR    22-OCT-1999;    99US-0160989.
PR    25-OCT-1999;    99US-0161404.
PR    25-OCT-1999;    99US-0161405.
PR    25-OCT-1999;    99US-0161406.
PR    26-OCT-1999;    99US-0161359.
PR    26-OCT-1999;    99US-0161360.
PR    26-OCT-1999;    99US-0161361.
PR    28-OCT-1999;    99US-0161920.
PR    28-OCT-1999;    99US-0161992.
PR    28-OCT-1999;    99US-0161993.
PR    29-OCT-1999;    99US-0162142.
```

```
   Query Match              10.1%;   Score 43.8;   DB 21;   Length 1519;
   Best Local Similarity    49.8%;   Pred. No. 0.0023;
   Matches  108;  Conservative    0;  Mismatches  109;  Indels    0;  Gaps    0;
```

```
Qy     64 gctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcgcgctgcggac 123
          ||| | | || || | || | || || || || || | |
Db   1297 gcttctggcgtttggattaggtagaagagcgtgtcctggatcgggtctggcccaacgaat 1356

Qy    124 cgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacggttgatggagc 183
          ||| || || | |||| | || | | |||||| || |||||| | |||| |
Db   1357 cgtgggactagctctcgggtcattgatacaatgctttgaatgggagagagttgggaatgt 1416

Qy    184 tcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcccgttggaggc 243
           | ||        | |     | ||| || | ||| || | || ||| | ||
Db   1417 ggaagtggatatgaaggaaggagttgggaatactgtacccaaagcgattcctttgaaagc 1476

Qy    244 catgtgcangccgcgtacagctatgcgtggtgttctt 280
          || |||| | ||| || | | | | || ||
Db   1477 tatttgcaaagctcgtccatttctacataagattatt 1513
```

Search completed: February 7, 2002, 11:00:07
Job time: 4993 sec

                          GenCore version 4.5
                Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2002, 11:12:10 ; Search time 172.96 Seconds
                                               (without alignments)
                                               565.671 Million cell updates/sec

Title: US-09-394-745-6514
Perfect score: 432
Sequence: 1 gtccagcagctcggacttac.........attttctttttttttcttgg 432

Scoring table: IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
                 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

         Pred. No. is the number of results predicted by chance to have a
         score greater than or equal to the score of the result being printed,
         and is derived by analysis of the total score distribution.

                                    SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	59.6	13.8	622	2	US-08-975-316-17	Sequence 17, Appl
	2	42	9.7	1506	4	US-09-158-767-7	Sequence 7, Appli
	3	42	9.7	1506	4	US-09-158-767-8	Sequence 8, Appli
	4	42	9.7	1506	4	US-09-158-767-9	Sequence 9, Appli
	5	42	9.7	2261	4	US-09-158-767-1	Sequence 1, Appli
	6	38.4	8.9	1762	3	US-08-881-784-5	Sequence 5, Appli
	7	38.4	8.9	1762	4	US-09-292-768-1	Sequence 1, Appli
	8	38.4	8.9	1762	4	US-09-292-768-63	Sequence 63, Appl
	9	38.4	8.9	1762	4	US-09-292-768-65	Sequence 65, Appl
	10	38.4	8.9	1762	4	US-09-172-339-5	Sequence 5, Appli
	11	37	8.6	1929	4	US-09-380-420C-1	Sequence 1, Appli
c	12	36	8.3	8438	1	US-07-945-283-1	Sequence 1, Appli
c	13	35.4	8.2	6387	1	US-07-721-775A-1	Sequence 1, Appli
c	14	35.4	8.2	6387	1	US-08-339-658-1	Sequence 1, Appli
	15	34.6	8.0	1219	4	US-09-025-819-28	Sequence 28, Appl
	16	34.6	8.0	11220	4	US-09-105-537-32	Sequence 32, Appl
	17	34.6	8.0	36778	4	US-09-105-537-5	Sequence 5, Appli
	18	34.6	8.0	38506	3	US-09-320-878-19	Sequence 19, Appl
	19	33.6	7.8	996	4	US-09-025-819-1	Sequence 1, Appli
	20	33.6	7.8	1515	4	US-09-292-768-5	Sequence 5, Appli
	21	33.6	7.8	1665	3	US-08-881-784-8	Sequence 8, Appli
	22	33.6	7.8	1665	4	US-09-292-768-3	Sequence 3, Appli
	23	33.6	7.8	1665	4	US-09-292-768-67	Sequence 67, Appl
	24	33.6	7.8	1665	4	US-09-292-768-69	Sequence 69, Appl
	25	33	7.6	1893	1	US-08-532-065B-1	Sequence 1, Appli
c	26	33	7.6	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	27	33	7.6	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	28	32.6	7.5	43280	2	US-08-804-227C-1	Sequence 1, Appli
c	29	32.2	7.5	1656	4	US-09-385-028-14	Sequence 14, Appl
c	30	32.2	7.5	15079	4	US-09-385-028-1	Sequence 1, Appli
	31	32	7.4	801	2	US-08-975-316-50	Sequence 50, Appl

```
    32   31.8   7.4   1518   1   US-08-148-215A-3        Sequence 3, Appli
    33   31.4   7.3 4403765   4   US-09-103-840A-2        Sequence 2, Appli
    34   31.4   7.3 4411529   4   US-09-103-840A-1        Sequence 1, Appli
 c  35   31.2   7.2    461   2   US-08-825-556A-1        Sequence 1, Appli
    36   31.2   7.2   1269   1   US-08-396-218-1         Sequence 1, Appli
    37   31.2   7.2   1269   1   US-08-760-116-1         Sequence 1, Appli
    38   30.4   7.0   1162   2   US-08-726-306A-52       Sequence 52, Appl
    39   30.4   7.0   1575   2   US-08-811-897A-34       Sequence 34, Appl
    40   30.4   7.0   1575   2   US-08-855-213-34        Sequence 34, Appl
    41   30.4   7.0   1656   2   US-08-811-897A-36       Sequence 36, Appl
    42   30.4   7.0   1656   2   US-08-855-213-36        Sequence 36, Appl
    43   30.4   7.0   1659   2   US-08-811-897A-35       Sequence 35, Appl
    44   30.4   7.0   1659   2   US-08-811-897A-37       Sequence 37, Appl
    45   30.4   7.0   1659   2   US-08-855-213-35        Sequence 35, Appl
```

ALIGNMENTS

RESULT   1
US-08-975-316-17
; Sequence 17, Application US/08975316
; Patent No. 5952486
;   GENERAL INFORMATION:
;   APPLICANT: BLOKSBERG, Leonard N., HAVUKKALA, Ilkka
;   APPLICANT: and GRIERSON, Alastair W.
;   TITLE OF INVENTION: MATERIALS AND METHODS FOR
;   TITLE OF INVENTION: THE MODIFICATION OF PLANT LIGNIN CONTENT
;   NUMBER OF SEQUENCES: 88
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Law Offices of Ann W. Speckman
;   STREET: 2601 Elliott Avenue, Suite 4185
;   CITY: Seattle
;   STATE: WA
;   COUNTRY: USA
;   ZIP: 98121
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/975,316
;   FILING DATE:
;   CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/713,000
;   FILING DATE: September 11, 1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: SLEATH, Janet
;   REGISTRATION NUMBER: 37,007
;   REFERENCE/DOCKET NUMBER: 11000/1003C1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 206-269-0565
;   TELEFAX: 206-269-0563
;   TELEX:
;   INFORMATION FOR SEQ ID NO: 17:
```

   Query Match                13.8%;   Score 59.6;   DB 2;   Length 622;
   Best Local Similarity     54.1%;   Pred. No. 1e-08;
   Matches  119;  Conservative    0;  Mismatches  101;  Indels    0;  Gaps    0;

Qy      61 cctgctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcgcgctgcg 120
            ||   ||   | || ||  ||||||||||| ||    | ||| || || |      || ||       |
Db     180 CCGACTATTGCCGTTTGGGATGGGGAGGAGAAGTTGTCCTGGTGCTGGCCTTGCCAATAG 239

Qy     121 gaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacggttgatgg 180
            ||   |   ||||  || ||   |||| |   |  | |||||||||| |||||     |||| |
Db     240 AGTGGTGAGCTTGGTCCTGGCGGCGCTTATTCAGTGCTTCGAATGGGAACGAGTTGGCGA 299

Qy     181 agctcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcccgttgga 240
            ||      |||      |           || || || ||||| |     |     || ||||||
Db     300 AGAATTGGTGGACTTGTCCGAGGGGACGGGACTCACAATGCCAAAGAGAGAGCCATTGGA 359

Qy     241 ggccatgtgcangccgcgtacagctatgcgtggtgttctt 280
            |||| |||||| ||||| |   ||| | |   |  | ||
Db     360 GGCCTTGTGCAAAGCGCGTGAATGCATGATAGCTAATGTT 399

   Query Match                 9.7%;   Score 42;   DB 4;   Length 1506;

```
  Best Local Similarity   53.7%;   Pred. No. 0.0027;
  Matches   87;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;


Qy      1 gtccagcagctcggacttacccggccggttcgatggctccggcggcaaggccaagggggcc 60
            || |||        |     || | ||      | | |   || | | | | ||| |
Db   1233 gttcaggccggagaggttcctcgaggaggagaaggccgtcgaggcccacggcaacgattt 1292


Qy     61 cctgctgatccctttcgggatggggcggcccaattgccccggggaaacgctcgcgctgcg 120
            || | |   | || ||||| | || || |   |  ||||||||||   | |||||||||||
Db   1293 ccggttcgtgcccttcggcgtcggccgccggagctgccccgggatcatcctcgcgctgcc 1352


Qy    121 gaccgtcgggctggtgctcgcaacgctgctcaattgcttcga 162
            | | |||| |    ||||| |   ||| | |    |||| |
Db   1353 catcatcggcatcacgctcggacgcctggtgcagaacttcca 1394
```

RESULT   3
US-09-158-767-8
; Sequence 8, Application US/09158767A
; Patent No. 6180363
; GENERAL INFORMATION:
; APPLICANT: Batard, Yannick
; APPLICANT: Durst, Francis
; APPLICANT: Schalk, Michel
; APPLICANT: Werck-Reichhart, Daniele
; TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING
; TITLE OF INVENTION: EXPRESSION IN YEAST AND OBTAINED TRANSFORMED YEAST
; FILE REFERENCE: A32000
; CURRENT APPLICATION NUMBER: US/09/158,767A
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: FR 97-12094
; EARLIER FILING DATE: 1997-09-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Altered sequences
US-09-158-767-8


```
  Query Match              9.7%;   Score 42;  DB 4;  Length 1506;
  Best Local Similarity   53.7%;   Pred. No. 0.0027;
  Matches   87;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;


Qy      1 gtccagcagctcggacttacccggccggttcgatggctccggcggcaaggccaagggggcc 60
            || |||        |     || | ||      | | |   || | | | | ||| |
Db   1233 gttcaggccggagaggttcctcgaggaggagaaggccgtcgaggcccacggcaacgattt 1292


Qy     61 cctgctgatccctttcgggatggggcggcccaattgccccggggaaacgctcgcgctgcg 120
            || | |   | || ||||| | || || |   |  ||||||||||   | |||||||||||
Db   1293 ccggttcgtgcccttcggcgtcggccgccggagctgccccgggatcatcctcgcgctgcc 1352


Qy    121 gaccgtcgggctggtgctcgcaacgctgctcaattgcttcga 162
            | | |||| |    ||||| |   ||| | |    |||| |
```

Db   1353 catcatcggcatcacgctcggacgcctggtgcagaacttcca 1394

RESULT   4
US-09-158-767-9
; Sequence 9, Application US/09158767A
; Patent No. 6180363
; GENERAL INFORMATION:
; APPLICANT: Batard, Yannick
; APPLICANT: Durst, Francis
; APPLICANT: Schalk, Michel
; APPLICANT: Werck-Reichhart, Daniele
; TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING
; TITLE OF INVENTION: EXPRESSION IN YEAST AND OBTAINED TRANSFORMED YEAST
; FILE REFERENCE: A32000
; CURRENT APPLICATION NUMBER: US/09/158,767A
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: FR 97-12094
; EARLIER FILING DATE: 1997-09-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 1506
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Altered sequences
US-09-158-767-9

  Query Match   9.7%; Score 42; DB 4; Length 1506;
  Best Local Similarity   53.7%; Pred. No. 0.0027;
  Matches 87; Conservative   0; Mismatches   75; Indels   0; Gaps   0;

Qy   1 gtccagcagctcggacttacccggccggttcgatggctccggcggcaaggccaagggggcc 60
            || ||| | || | || | | | | || | | | | ||| |
Db   1233 gttcaggccggagaggttcctcgaggaggagaaggccgtcgaggcccacggcaacgattt 1292

Qy   61 cctgctgatccctttcgggatggggcggcccaattgccccggggaaacgctcgcgctgcg 120
            || | | | || ||||| | || || | | ||||||||| | |||||||||||
Db   1293 ccggttcgtgcccttcggcgtcggccgccggagctgccccgggatcatcctcgcgctgcc 1352

Qy   121 gaccgtcgggctggtgctcgcaacgctgctcaattgcttcga 162
            | | |||| | |||||| | ||| | | |||| |
Db   1353 catcatcggcatcacgctcggacgcctggtgcagaacttcca 1394

RESULT   5
US-09-158-767-1
; Sequence 1, Application US/09158767A
; Patent No. 6180363
; GENERAL INFORMATION:
; APPLICANT: Batard, Yannick
; APPLICANT: Durst, Francis
; APPLICANT: Schalk, Michel
; APPLICANT: Werck-Reichhart, Daniele
; TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING

  Query Match   9.7%; Score 42; DB 4; Length 2261;
  Best Local Similarity   53.7%; Pred. No. 0.0032;
  Matches   87; Conservative   0; Mismatches   75; Indels   0; Gaps   0;

Qy   1 gtccagcagctcggacttacccggccggttcgatggctccggcggcaaggccaagggggcc 60
          || ||| | || | || | ||| | || | | | | ||| |
Db   1281 gttcaggccggagaggttcctcgaggaggagaaggccgtcgaggcccacggcaacgattt 1340

Qy   61 cctgctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcgcgctgcg 120
          || | | | || |||||| | || || | | |||||||||| | ||||||||||
Db   1341 ccggttcgtgcccttcggcgtcggccgccggagctgccccgggatcatcctcgcgctgcc 1400

Qy   121 gaccgtcgggctggtgctcgcaacgctgctcaattgcttcga 162
           | | |||| | ||||| | ||| | | |||| |
Db   1401 catcatcggcatcacgctcggacgcctggtgcagaacttcca 1442

RESULT   6
US-08-881-784-5
; Sequence 5, Application US/08881784
; Patent No. 6083731
;   GENERAL INFORMATION:
;   APPLICANT: Croteau, Rodney B.
;   APPLICANT: Lupien, Shari L.
;   APPLICANT: Karp, Frank
;   TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
;   TITLE OF INVENTION: THE PRODUCTION OF LIMONENE HYDROXYLASES
;   NUMBER OF SEQUENCES: 58
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
;   ADDRESSEE: PLLC
;   STREET: 1420 Fifth Avenue, Suite 2800
;   CITY: Seattle
;   STATE: WA
;   COUNTRY: USA
;   ZIP: 98101
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:

; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1762 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   ORIGINAL SOURCE:
;   ORGANISM: Mentha spicata
;   INDIVIDUAL ISOLATE: cDNA encoding
;   INDIVIDUAL ISOLATE:
;   IMMEDIATE SOURCE:
;   CLONE: pSM12.2
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 558..1212
;   OTHER INFORMATION: /product= "Probe LH-1 (Figure 4A)"
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 39..538
;   OTHER INFORMATION: /product= "Probe LH-2 (Figure 4A)"
US-08-881-784-5

   Query Match   8.9%;   Score 38.4;   DB 3;   Length 1762;
   Best Local Similarity   55.1%;   Pred. No. 0.034;
   Matches   75; Conservative   0; Mismatches   61; Indels   0; Gaps   0;

Qy   46 caaggccaaggggcccctgctgatcccctttcgggatggggcggcccaattgccccggggga 105
           || || || | | | | |||||| |||||| ||| || | ||||||||
Db   1276 CATGGGAAACGATTTCGAGTTCATCCCATTCGGGGCGGGTCGAAGAATCTGCCCCGGTTT 1335

Qy   106 aacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactg 165
           | ||| |||| | || | | | || ||||| | |||||||||||||
Db   1336 ACATTTCGGGCTGGCAAATGTTGAGATCCCATTGGCGCAACTGCTCTACCACTTCGACTG 1395

Qy   166 ggacacggttgatgga 181
           | | | | |||
Db   1396 GAAATTGCCACAAGGA 1411

RESULT   7
US-09-292-768-1
; Sequence 1, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B

   Query Match   8.9%;   Score 38.4;   DB 4;   Length 1762;
   Best Local Similarity   55.1%;   Pred. No. 0.034;
   Matches   75; Conservative   0; Mismatches 61; Indels   0; Gaps   0;

Qy   46 caaggccaaggggcccctgctgatccctttcgggatggggcggcccaattgccccggggga 105
           || || || | | | | |||||| |||||| ||| || | ||||||||
Db   1276 catgggaaacgatttcgagttcatcccattcggggcgggtcgaagaatctgccccggtttt 1335

Qy   106 aacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactg 165
           | ||| |||| | || | | | | || |||||| | ||||||||||
Db   1336 acatttcgggctggcaaatgttgagatcccattggcgcaactgctctaccacttcgactg 1395

Qy   166 ggacacggttgatgga 181
            | | | | | |||
Db   1396 gaaattgccacaagga 1411

  Query Match   8.9%;   Score 38.4;   DB 4;   Length 1762;
  Best Local Similarity   55.1%;   Pred. No. 0.034;
  Matches   75; Conservative   0; Mismatches   61; Indels   0; Gaps   0;

Qy   46 caaggccaagggggcccctgctgatccctttcgggatggggcggcccaattgccccggggga 105
           || || || | | | | |||||| |||||| ||| || | |||||||||
Db   1276 catgggaaacgatttcgagttcatcccattcggggcgggtcgaagaatctgccccggttt 1335

Qy   106 aacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactg 165
           | ||| |||| | || | | | | || |||||| | |||||||||||
Db   1336 acatttcgggctggcaaatgttgagatcccattggcgcaactgctctaccacttcgactg 1395

Qy   166 ggacacggttgatgga 181
            | | | | |||
Db   1396 gaaattgccacaagga 1411

RESULT   9
US-09-292-768-65
; Sequence 65, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; TITLE OF INVENTION: LIMONENE HYDROXYLASES
; FILE REFERENCE: wsur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; EARLIER FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 65
;   LENGTH: 1762
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: computer-generated nucleic acid sequence encoding
;   OTHER INFORMATION: limonene-6-hydroxylase
;   FEATURE:
;   NAME/KEY: misc_feature

```
;   LOCATION: (1)..(1762)
;   OTHER INFORMATION: computer-generated nucleic acid sequence encoding
;   OTHER INFORMATION: spearmint limonene-6-hydroxylase variant
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (20)..(1507)
US-09-292-768-65


   Query Match               8.9%;   Score 38.4;   DB 4;   Length 1762;
   Best Local Similarity    55.1%;   Pred. No. 0.034;
   Matches   75; Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy     46 caaggccaaggggcccctgctgatccctttcgggatggggcggcccaattgccccggggga 105
          || ||   || |       |   | | ||||| |||||| ||| ||    |  |||||||||
Db   1276 catgggaaacgatttcgagttcatcccattcggggcgggtcgaagaatctgccccggtttt 1335

Qy    106 aacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactg 165
          |    ||| ||||  |  ||  | | |   | ||     |||||| |   |||||||||||
Db   1336 acatttcgggctggcaaatgttgagatcccattggcgcaactgctctaccacttcgactg 1395

Qy    166 ggacacggttgatgga 181
          | |   |      | |||
Db   1396 gaaattgccacaagga 1411


RESULT  10
US-09-172-339-5
; Sequence 5, Application US/09172339
; Patent No. 6291745
; GENERAL INFORMATION:
;   APPLICANT: EuClaire Meyer, Terry
;   APPLICANT: Yalpani, Nasser
;   TITLE OF INVENTION: Limonene and Other Downstream
;   TITLE OF INVENTION: Metabolites of Geranyl Pyrophosphate for Insect Control
in
;   TITLE OF INVENTION: Plants
;   FILE REFERENCE: 5718-65
;   CURRENT APPLICATION NUMBER: US/09/172,339
;   CURRENT FILING DATE: 1998-10-14
;   EARLIER APPLICATION NUMBER: 08/449,061
;   EARLIER FILING DATE: 1995-05-24
;   EARLIER APPLICATION NUMBER: 08/153,544
;   EARLIER FILING DATE: 1993-11-16
;   EARLIER APPLICATION NUMBER: 08/042,199
;   EARLIER FILING DATE: 1993-04-02
;   NUMBER OF SEQ ID NOS: 8
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 1762
;   TYPE: DNA
;   ORGANISM: Mentha spicata
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Carveol Synthase
;   FEATURE:
```

```
   Query Match              8.9%;   Score 38.4;   DB 4;   Length 1762;
   Best Local Similarity   55.1%;   Pred. No. 0.034;
   Matches    75;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;


Qy     46 caaggccaaggggcccctgctgatccctttcgggatggggcggcccaattgccccgggga 105
          || ||  || |     |  | | |||||| |||||| ||| ||    |  ||||||||
Db   1276 catgggaaacgatttcgagttcatcccattcggggcgggtcgaagaatctgccccggttt 1335


Qy    106 aacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactg 165
          |    ||| ||||   |  || | | |       | ||    |||||| |   ||||||||||
Db   1336 acatttcgggctggcaaatgttgagatcccattggcgcaactgctctaccacttcgactg 1395


Qy    166 ggacacggttgatgga 181
          | |    |    | |||
Db   1396 gaaattgccacaagga 1411
```


RESULT 11
US-09-380-420C-1
; Sequence 1, Application US/09380420C
; Patent No. 6300544
;   GENERAL INFORMATION:
;   APPLICANT: Halkier, Barbara
;   Bak, Soren
;   Kahn, Rachel
;   Moller, Birger
;   TITLE OF INVENTION: Cytochrome P450 Monooxygenases
;   NUMBER OF SEQUENCES: 23
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Syngenta Patent Dept.
;   STREET: 3054 Cornwallis Road
;   CITY: RTP
;   STATE: NC
;   COUNTRY: USA
;   ZIP: 27709
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/380,420C
;   FILING DATE: 12-No. 6300544-1999
;   CLASSIFICATION: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Meigs, J. Timothy
;   REGISTRATION NUMBER: 38,241
;   REFERENCE/DOCKET NUMBER: S-21251A
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 919-541-8587
;   INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:

   Query Match   8.6%;   Score 37;   DB 4;   Length 1929;
   Best Local Similarity   54.9%;   Pred. No. 0.092;
   Matches   73; Conservative   0; Mismatches   60; Indels   0; Gaps   0;

Qy   35 ggctccggcggcaaggccaagggggcccctgctgatccctttcgggatggggcggcccaat 94
          | | || | | | | | | | ||| || || |||||| || || | |
Db   1422 GACGTCGACTACTACGGCTCGCACTTCGAGCTCATACCGTTCGGGGCCGGCCGCCGGATC 1481

Qy   95 tgccccggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaat 154
          ||||| || || | | | | |||| |||||| | ||||||| |
Db   1482 TGCCCGGGACTCACCATGGGCGAGACCAACGTCACCTTCACCCTCGCCAACCTGCTCTAC 1541

Qy   155 tgcttcgactggg 167
          |||| ||||||||
Db   1542 TGCTACGACTGGG 1554

RESULT 12
US-07-945-283-1/c
; Sequence 1, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
;   APPLICANT: Cheung, Andrew K.
;   APPLICANT: Wesley, Ronald D.
;   TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
;   TITLE OF INVENTION: Involving The EP0 and LLT Genes
;   NUMBER OF SEQUENCES: 7
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Curtis P. Ribando
;   STREET: 1815 No. 5352596th University Street
;   CITY: Peoria
;   STATE: IL
;   COUNTRY: USA
;   ZIP: 61604
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;.   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/945,283
;   FILING DATE: 19920911
;   CLASSIFICATION: 424

; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 8438 base pairs
;   TYPE: NUCLEIC ACID
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;   ORGANISM: Pseudorabies virus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 622..6495
;   FEATURE:
;   NAME/KEY: variation
;   LOCATION: replace(1099, "g")
;   FEATURE:
;   NAME/KEY: variation
;   LOCATION: replace(1267, "t")
;   FEATURE:
;   NAME/KEY: variation
;   LOCATION: replace(1381, "c")
;   FEATURE:
;   NAME/KEY: variation
;   LOCATION: replace(1566, "c")
;   FEATURE:
;   NAME/KEY: variation
;   LOCATION: replace(7010, "g")
US-07-945-283-1

   Query Match   8.3%;   Score 36;   DB 1;   Length 8438;
   Best Local Similarity   50.0%;   Pred. No. 0.34;
   Matches   90;   Conservative   0;   Mismatches   90;   Indels   0;   Gaps   0;

Qy   22 cggccggttcgatggctccggcggcaaggccaagggggcccctgctgatccctttcgggat 81
          || | | || | ||| | ||||| | || | | |||| ||| || || |
Db   4708 CGCCTGCGTCCTGGCCTGCCGCGGCGTCCTCGAGCGCCTGCTGCCCTGCCCGCTCCGGCT 4649

Qy   82 ggggcggcccaattgccccgggggaaacgctcgcgctgcggaccgtcgggctggtgctcgc 141
          | |||| || || || | |||| || | ||| | |||| |||
Db   4648 GCCCGCGCCCGCCCGCGCCCCGGCCGCCCTCGGGCCCGCCTGCCTCGAGGAGGTGACCGC 4589

Qy   142 aacgctgctcaattgcttcgactgggacacggttgatggagctcaggtttgacatgaagc 201
          |||||||| | |||| | | | | | || | | | || | ||
Db   4588 CGCGCTGCTCGCGCTCCGCGACGCGATCCCCGGGGCCGGCCCGGCCGAGCGGCAGCAGGC 4529

RESULT   13
US-07-721-775A-1/c

```
;  Sequence 1, Application US/07721775A
;  Patent No. 5180666
;   GENERAL INFORMATION:
;      APPLICANT:  States, J. Christopher
;      APPLICANT:  Hines, Ronald N.
;      APPLICANT:  No. 5180666ak, Raymond F.
;      TITLE OF INVENTION:  METHOD AND CELL LINE FOR TESTING
;      TITLE OF INVENTION:  MUTAGENICITY OF A CHEMICAL
;      NUMBER OF SEQUENCES:  2
;      CORRESPONDENCE ADDRESS:
;        ADDRESSEE:  Reising, Ethington, Barnard, Perry & Milton
;        STREET:  P.O. Box 4390
;        CITY:  Troy
;        STATE:  Michigan
;        COUNTRY:  U.S.A.
;        ZIP:  48099
;      COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Floppy disk
;        COMPUTER:  IBM PC compatible
;        OPERATING SYSTEM:  PC-DOS/MS-DOS
;        SOFTWARE:  PatentIn Release #1.0, Version #1.25
;      CURRENT APPLICATION DATA:
;        APPLICATION NUMBER:  US/07/721,775A
;        FILING DATE:  19910627
;        CLASSIFICATION:  435
;      ATTORNEY/AGENT INFORMATION:
;        NAME:  Kohn, Kenneth I.
;        REGISTRATION NUMBER:  30,955
;        REFERENCE/DOCKET NUMBER:  P-321WSU
;      TELECOMMUNICATION INFORMATION:
;        TELEPHONE:  (313) 689-3554
;        TELEFAX:  (313) 689-4071
;   INFORMATION FOR SEQ ID NO:  1:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH:  6387 base pairs
;        TYPE:  NUCLEIC ACID
;        STRANDEDNESS:  double
;        TOPOLOGY:  circular
;      MOLECULE TYPE:  DNA (genomic)
;      ORIGINAL SOURCE:
;        ORGANISM:  Homo sapiens
;      FEATURE:
;        NAME/KEY:  exon
;        LOCATION:  complement (2807..3631)
;      FEATURE:
;        NAME/KEY:  exon
;        LOCATION:  complement (2125..2251)
;      FEATURE:
;        NAME/KEY:  exon
;        LOCATION:  complement (1948..2037)
;      FEATURE:
;        NAME/KEY:  exon
;        LOCATION:  complement (1733..1856)
;      FEATURE:
;        NAME/KEY:  exon
;        LOCATION:  complement (1501..1587)
;      FEATURE:
```

```
;       NAME/KEY:  exon
;       LOCATION:  complement (237..1308)
;    FEATURE:
;       NAME/KEY:  promoter
;       LOCATION:  complement (3638..3967)
;    FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  4586..5446
US-07-721-775A-1


   Query Match              8.2%;   Score 35.4;   DB 1;   Length 6387;
   Best Local Similarity   48.5%;   Pred. No. 0.46;
   Matches   96;  Conservative   0;  Mismatches 102;  Indels   0;  Gaps   0;

Qy      32 gatggctccggcggcaaggccaaggggcccctgctgatccctttcgggatggggcggccc 91
           |||||  |   || ||||||        |    | ||||   || || ||||||   ||
Db    1256 GATGGTGCTATCGACAAGGTGTTAAGTGAGAAGGTGATTATCTTTGGCATGGGCAAGCGG 1197

Qy      92 aattgccccgggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctc 151
           || ||   ||| || || | || |  || |   ||| || | || || |  |||||
Db    1196 AAGTGTATCGGTGAGACCATTGCCCGCTGGGAGGTCTTTCTCTTCCTGGCTATCCTGCTG 1137

Qy     152 aattgcttcgactgggacacggttgatggagctcaggtttgacatgaagctancggcggg 211
            |  |  | || |    | |        || |   ||||          | | |   |||
Db    1136 CAACGGGTGGAATTCAGCGTGCCACTGGGCGTGAAGGTGGACATGACCCCCATCTATGGG 1077

Qy     212 ctgaccatgccccgggcc 229
           || |||||||   |  |||
Db    1076 CTAACCATGAAGCATGCC 1059



RESULT  14
US-08-339-658-1/c
; Sequence 1, Application US/08339658
; Patent No. 5525482
;   GENERAL INFORMATION:
;     APPLICANT:  States, J. Christopher
;     APPLICANT:  Hines, Ronald N.
;     APPLICANT:  No. 5525482ak, Raymond F.
;     TITLE OF INVENTION:  METHOD AND CELL LINE FOR TESTING
;     TITLE OF INVENTION:  MUTAGENICITY OF A CHEMICAL
;     NUMBER OF SEQUENCES:  2
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Reising, Ethington, Barnard, Perry & Milton
;       STREET:  P.O. Box 4390
;       CITY:  Troy
;       STATE:  Michigan
;       COUNTRY:  U.S.A.
;       ZIP:  48099
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/339,658
```

; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 6387 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   ORIGINAL SOURCE:
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: complement (2807..3631)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: complement (2125..2251)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: complement (1948..2037)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: complement (1733..1856)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: complement (1501..1587)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: complement (237..1308)
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: complement (3638..3967)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 4586..5446
US-08-339-658-1

   Query Match   8.2%;   Score 35.4; DB 1; Length 6387;
   Best Local Similarity   48.5%; Pred. No. 0.46;
   Matches   96; Conservative   0; Mismatches 102; Indels   0; Gaps   0;

Qy   32 gatggctccggcggcaaggccaaggggcccctgctgatccctttcgggatggggcggccc 91
          ||||| | || ||||| | | |||| || || ||||| ||
Db   1256 GATGGTGCTATCGACAAGGTGTTAAGTGAGAAGGTGATTATCTTTGGCATGGGCAAGCGG 1197

Qy   92 aattgccccgggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctc 151

```
          ||  ||    ||| || ||   | || |    ||    |||    ||  | || || |  |||||
Db   1196 AAGTGTATCGGTGAGACCATTGCCCGCTGGGAGGTCTTTCTCTTCCTGGCTATCCTGCTG 1137

Qy    152 aattgcttcgactgggacacggttgatggagctcaggtttgacatgaagctancggcggg 211
              |  |  | ||  |     |  |      || |   ||||       | | |    |||
Db   1136 CAACGGGTGGAATTCAGCGTGCCACTGGGCGTGAAGGTGGACATGACCCCCATCTATGGG 1077

Qy    212 ctgaccatgccccgggcc 229
              ||  ||||||  |  |||
Db   1076 CTAACCATGAAGCATGCC 1059


RESULT  15
US-09-025-819-28
; Sequence 28, Application US/09025819
; Patent No. 6225097
;  GENERAL INFORMATION:
;     APPLICANT:  Obata, Shusei
;     APPLICANT:  Nishino, Tokuzo
;     APPLICANT:  Koyama, Tanetoshi
;     APPLICANT:  Sato, Yoshihiro
;     TITLE OF INVENTION:  DECAPRENYL DIPHOSPHATE SYNTHETASE GENE
;     NUMBER OF SEQUENCES:  31
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  KENYON & KENYON
;       STREET:  1500 K Street, N.W.
;       CITY:  Washington
;       STATE:  DC
;       COUNTRY:  USA
;       ZIP:  20005
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/025,819
;       FILING DATE:  19-FEB-1998
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP 251675
;       FILING DATE:  17-SEP-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Khalilian, Houri
;       REGISTRATION NUMBER:  39,546
;       REFERENCE/DOCKET NUMBER:  10235/2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-220-4200
;       TELEFAX:  202-220-4201
;  INFORMATION FOR SEQ ID NO:  28:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1219 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
```

   Query Match   8.0%;   Score 34.6;   DB 4;   Length 1219;
   Best Local Similarity   53.3%;   Pred. No. 0.4;
   Matches   73; Conservative   0; Mismatches   64; Indels   0; Gaps   0;

Qy   3 ccagcagctcggacttacccggccggttcgatggctccggcggcaaggccaagggggcccc 62
          || ||| |||| | ||| | | | | ||||||||||| | |·|||
Db   266 CCCGCATTCCGGAAGTGACCGCGCATCTGGTCGAGGCCGGCGGCAAGCGGCTGCGGCCGA 325

Qy   63 tgctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcgcgctgcgga 122
          ||||| || | | | | ||| | | | ||| | | | | |||||| |
Db   326 TGCTGGTGCTGGCGGCGGCGCGGCTGTGCGGCTATCAGGGGAACAGCCATGTGCTGCTGG 385

Qy   123 ccgtcgggctggtgctc 139
          ||| | | | | | ||
Db   386 CCGCGGCGGTCGAGTTC 402

Search completed: February 7, 2002, 11:12:24
Job time: 7310 sec

                        GenCore version 4.5
                Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2002, 08:20:51 ; Search time 4942.22 Seconds
                                              (without alignments)
                                              939.290 Million cell updates/sec

Title: US-09-394-745-6514
Perfect score: 432
Sequence: 1 gtccagcagctcggacttac..........attttctttttttttcttgg 432

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
                 1: em_estfun:*
                 2: em_esthum:*

```
 3:  em_estin:*
 4:  em_estom:*
 5:  em_estpl:*
 6:  em_estba:*
 7:  em_estro:*
 8:  em_estov:*
 9:  em_htc:*
10:  gb_est1:*
11:  gb_est2:*
12:  gb_htc:*
13:  gb_gss:*
14:  em_gss_fun:*
15:  em_gss_hum:*
16:  em_gss_inv:*
17:  em_gss_pln:*
18:  em_gss_pro:*
19:  em_gss_rod:*
20:  em_gss_vrt:*
21:  em_gss_other:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	179	41.4	430	10	AW922538	AW922538 DG1_20_D1
	2	179	41.4	501	10	BE363286	BE363286 WS1_61_D0
	3	179	41.4	594	10	BE355191	BE355191 DG1_10_D0
	4	179	41.4	634	10	BE360028	BE360028 DG1_60_C0
	5	179	41.4	654	10	BE362029.	BE362029 DG1_83_E0
	6	179	41.4	693	10	AW676742	AW676742 DG1_14_A0
	7	179	41.4	695	10	BE357860	BE357860 DG1_22_C1
	8	169.2	39.2	535	10	AW922289	AW922289 DG1_17_H0
c	9	163	37.7	584	10	AI668207	AI668207 605018C02
c	10	150.4	34.8	790	11	BG320973	BG320973 Zm04_02d0
	11	150.2	34.8	357	11	BG464759	BG464759 EM1_33_G0
	12	148.6	34.4	293	11	BG464902	BG464902 EM1_35_G0
c	13	145.8	33.8	803	10	BE704790	BE704790 Sc02_02e1
c	14	142.4	33.0	700	10	AL503532	AL503532 AL503532
c	15	142.4	33.0	702	10	BE412662	BE412662 MCG007.D1
	16	141.4	32.7	494	10	BE445503	BE445503 WHE1135_C
c	17	140.2	32.5	679	10	BE418633	BE418633 SCL072.F0
c	18	137.8	31.9	644	13	AQ288789	AQ288789 nbxb0033H
	19	131	30.3	268	11	BG464757	BG464757 EM1_33_G0
c	20	130.2	30.1	526	10	AI920363	AI920363 603020B03
c	21	130.2	30.1	620	10	AI947634	AI947634 603025F12
	22	123.8	28.7	719	11	C97610	C97610 C97610 Rice
c	23	111.4	25.8	421	10	AL502922	AL502922 AL502922
	24	110	25.5	556	10	AW679544	AW679544 WS1_29_D0
c	25	109.4	25.3	338	11	BG905630	BG905630 TaLr1141A
	26	97.4	22.5	932	10	AI001245	AI001245 MEST5-E8.
	27	81.2	18.8	449	10	AU173235	AU173235 AU173235
	28	79.2	18.3	516	10	AW285419	AW285419 LG1_251_B

```
        29    76.8   17.8    374   10   BE425382        BE425382 WHE313_G0
        30    72     16.7    589   10   AA753913        AA753913 97BS0361
        31    66.8   15.5    424   10   AU070733        AU070733 AU070733
        32    65.8   15.2    507   11   BG933110        BG933110 WS1_29_D0
        33    64.8   15.0    700   10   AL504503        AL504503 AL504503
        34    61.8   14.3    537   10   AW257915        AW257915 687064B01
   c    35    59     13.7    523   10   AV557146        AV557146 AV557146
        36    55     12.7    638   10   BE360873        BE360873 DG1_67_A0
        37    54.4   12.6    722   11   BI305808        BI305808 NL_1_N04
   c    38    53.2   12.3    552   10   AW499362        AW499362 660053D07
        39    52.2   12.1    266   10   AA597575        AA597575 29483 Lam
        40    52.2   12.1    553   10   BE359396        BE359396 DG1_40_F0
   c    41    52     12.0    592   10   AW775060        AW775060 EST334211
        42    51.6   11.9    413   10   AA754418        AA754418 97MJ0362
        43    51.6   11.9    587   10   BE364385        BE364385 PI1_13_B1
        44    51     11.8    555   10   AW927862        AW927862 945013H06
        45    51     11.8    743   11   BI305417        BI305417 NLP_1_A13
```

ALIGNMENTS

RESULT   1
AW922538
LOCUS   AW922538   430 bp   mRNA   EST   19-JUL-2000
DEFINITION DG1_20_D10.g1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   AW922538
VERSION   AW922538.1 GI:8088363
KEYWORDS   EST.
SOURCE   sorghum.
  ORGANISM Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1 (bases 1 to 430)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
            ,L.H.
  TITLE   An EST database from Sorghum: dark-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: PolyTMix
            High quality sequence start: 103
            High quality sequence stop: 430
            POLYA=No.
FEATURES   Location/Qualifiers
     source   1. .430
                     /organism="Sorghum bicolor"

```
                        /db_xref="taxon:4558"
                        /clone_lib="Dark Grown 1 (DG1)"
                        /note="Organ: 5-day-old dark-grown seedlings; Vector:
                        Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
                        made from poly-A RNA in the cloning vector lambda ZAP II.
                        Clones to be sequenced were prepared by mass excision."
BASE COUNT        76 a     106 c     155 g     93 t
ORIGIN
```

```
   Query Match              41.4%;   Score 179;  DB 10;   Length 430;
   Best Local Similarity    78.8%;   Pred. No. 2.4e-33;
   Matches 212;  Conservative    0;  Mismatches   57;  Indels   0;  Gaps    0;
```

```
Qy     41 ggcggcaaggccaaggggcccctgctgatccctttcgggatggggcggcccaattgcccc 100
          | |||||||||| ||||||  |||||||||| || |||||||||||||||||| ||| ||||||
Db     27 GACGGCAAGGCCGAGGGGCGGCTGCTGATGCCGTTCGGGATGGGGCGGCGCAAGTGCCCC 86

Qy    101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160
          ||||| |||||||||||||||||||||||||||||||||| |||||| || | ||| ||
Db     87 GGGGAGACGCTCGCGCTGCGGACCGTCGGGCTGGTGCTCGGCACGCTGATCCAGTGCATC 146

Qy    161 gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg 220
          |||||||||| || ||||||      || ||           ||      |||||||||||||||
Db    147 GACTGGGACAGAGTCGATGGCCTGGAGATTGACATGACCGCGGGTGGCGGGCTGACCATG 206

Qy    221 ccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatgcgtggtgttctt 280
          ||| ||||||'|||||||||||||||||| ||||| ||| ||| |||||||||| | ||||||
Db    207 CCCAGGGCCGTCCCGTTGGAGGCCACGTGCAAGCCTCGTGCAGCTATGCGCGATGTTCTG 266

Qy    281 aagaggctctgaaaacctcatggatcgaa 309
          | |  |||||||  || ||| |  | |
Db    267 ATGGAGCTCTGAGCCTCTGATGAAGAGTA 295
```

```
RESULT    2
BE363286
LOCUS        BE363286       501 bp     mRNA            EST        20-JUL-2000
DEFINITION   WS1_61_D09.g1_A002 Water-stressed 1 (WS1) Sorghum bicolor cDNA,
             mRNA sequence.
ACCESSION    BE363286
VERSION      BE363286.1  GI:9304843
KEYWORDS     EST.
SOURCE       sorghum.
  ORGANISM   Sorghum bicolor
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
             clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE    1  (bases 1 to 501)
  AUTHORS    Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
             ,L.H.
  TITLE      An EST database from Sorghum: water-stressed plants
  JOURNAL    Unpublished (2000)
COMMENT      Contact: Cordonnier-Pratt MM
             Department of Botany
             The University of Georgia
             Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
```

FEATURES   Location/Qualifiers
        source   1. .501
                         /organism="Sorghum bicolor"
                         /db_xref="taxon:4558"
                         /clone_lib="Water-stressed 1 (WS1)"
                         /note="Organ: Mix of 5-week old plants on days 7 & 8 after
                         water was withheld; Vector: Lambda Zap; Site_1: XhoI;
                         Site_2: EcoRI; The library was made from poly-A RNA in the
                         cloning vector lambda ZAP II. Clones to be sequenced were
                         prepared by mass excision."
BASE COUNT   82 a   134 c   190 g   95 t
ORIGIN

    Query Match   41.4%; Score 179; DB 10; Length 501;
    Best Local Similarity   78.8%; Pred. No. 2.4e-33;
    Matches 212; Conservative   0; Mismatches   57; Indels   0; Gaps   0;

Qy   41 ggcggcaaggccaagggggcccctgctgatccctttcgggatggggcggcccaattgcccc 100
           | ||||||||||| |||||| |||||||| || |||||||||||||||| ||| ||||||
Db   136 GACGGCAAGGCCGAGGGGCGGCTGCTGATGCCGTTCGGGATGGGGCGGCGCAAGTGCCCC 195

Qy   101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160
           |||||| ||||||||||||||||||||||||||||||||||||||| |||||| || | ||| ||
Db   196 GGGGAGACGCTCGCGCTGCGGACCGTCGGGCTGGTGCTCGGCACGCTGATCCAGTGCATC 255

Qy   161 gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg 220
           |||||||||| || ||||| || || || ||||||||||||||||
Db   256 GACTGGGACAGAGTCGATGGCCTGGAGATTGACATGACCGCGGGTGGCGGGCTGACCATG 315

Qy   221 ccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatgcgtggtgttctt 280
           ||| ||||||||||||||||||||||| ||||| ||| ||| |||||||||| | |||||||
Db   316 CCCAGGGCCGTCCCGTTGGAGGCCACGTGCAAGCCTCGTGCAGCTATGCGCGATGTTCTG 375

Qy   281 aagaggctctgaaaacctcatggatcgaa 309
           | | ||||||| || ||| | | ||
Db   376 ATGGAGCTCTGAGCCTCTGATGAAGAGTA 404

RESULT   3
BE355191
LOCUS   BE355191   594 bp   mRNA   EST   20-JUL-2000
DEFINITION DG1_10_D02.g1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   BE355191
VERSION   BE355191.1 GI:9296181
KEYWORDS   EST.

```
SOURCE        sorghum.
  ORGANISM    Sorghum bicolor
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
              clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE     1  (bases 1 to 594)
  AUTHORS     Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
              ,L.H.
  TITLE       An EST database from Sorghum: dark-grown seedlings
  JOURNAL     Unpublished (2000)
COMMENT       Contact: Cordonnier-Pratt MM
              Department of Botany
              The University of Georgia
              Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
              Tel: 706 542 1860
              Fax: 706 542 1805
              Email: mmpratt@uga.edu
              Sequences have been trimmed to exclude PolyA, vector and regions
              below Phred quality 16. The threshold for highest quality sequence
              is 20.
              Seq primer: PolyTMix
              High quality sequence start: 27
              High quality sequence stop: 543
              POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .594
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="Dark Grown 1 (DG1)"
                     /note="Organ: 5-day-old dark-grown seedlings; Vector:
                     Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
                     made from poly-A RNA in the cloning vector lambda ZAP II.
                     Clones to be sequenced were prepared by mass excision."
BASE COUNT        100 a     172 c     213 g     109 t
ORIGIN
```

```
  Query Match                41.4%;   Score 179;   DB 10;   Length 594;
  Best Local Similarity      78.8%;   Pred. No. 2.4e-33;
  Matches  212;  Conservative    0;  Mismatches    57;  Indels    0;  Gaps    0;


Qy     41 ggcggcaaggccaagggggcccctgctgatccctttcgggatggggcggcccaattgcccc 100
          | ||||||||| |||||| ||||||||| || ||||||||||||||||| ||| ||||||
Db    229 GACGGCAAGGCCGAGGGGCGGCTGCTGATGCCGTTCGGGATGGGGCGGCGCAAGTGCCCC 288

Qy    101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160
          ||||| ||||||||||||||||||||||||||||||||||||| ||||||| || | ||| ||
Db    289 GGGGAGACGCTCGCGCTGCGGACCGTCGGGCTGGTGCTCGGCACGCTGATCCAGTGCATC 348

Qy    161 gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg 220
          |||||||||| || ||||||     || ||        ||      |||||||||||||||||
Db    349 GACTGGGACAGAGTCGATGGCCTGGAGATTGACATGACCGCGGGTGGCGGGCTGACCATG 408

Qy    221 ccccgggccgtcccgttggaggccatgtgcagccgcgtacagctatgcgtggtgttctt 280
          ||| ||||||||||||||||||||||| ||||| ||| ||| |||||||||| | ||||||
Db    409 CCCAGGGCCGTCCCGTTGGAGGCCACGTGCAAGCCTCGTGCAGCTATGCGCGATGTTCTG 468
```

```
Qy      281  aagaggctctgaaaacctcatggatcgaa 309
             |  |   |||||||    || ||| |  | |
Db      469  ATGGAGCTCTGAGCCTCTGATGAAGAGTA 497


RESULT    4
BE360028
LOCUS       BE360028       634 bp    mRNA            EST        20-JUL-2000
DEFINITION  DG1_60_C08.g2_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
            .           sequence.
ACCESSION   BE360028
VERSION     BE360028.1  GI:9301585
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
            ,L.H.
  TITLE     An EST database from Sorghum: dark-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: PolyTMix
            High quality sequence start: 38
            High quality sequence stop: 629
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .634
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="Dark Grown 1 (DG1)"
                     /note="Organ: 5-day-old dark-grown seedlings; Vector:
                     Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
                     made from poly-A RNA in the cloning vector lambda ZAP II.
                     Clones to be sequenced were prepared by mass excision."
BASE COUNT      116 a     153 c     219 g     146 t
ORIGIN


    Query Match              41.4%;   Score 179;  DB 10;   Length 634;
    Best Local Similarity    78.8%;   Pred. No. 2.4e-33;
    Matches 212; Conservative   0;  Mismatches   57;  Indels   0;  Gaps   0;


Qy      41  ggcggcaaggccaagggggcccctgctgatccctttcgggatggggcggcccaattgcccc 100
            | |||||||||| ||||||  |||||||| || ||||||||||||||||| ||| ||||||
Db     136  GACGGCAAGGCCGAGGGGCGGCTGCTGATGCCGTTCGGGATGGGGCGGCGCAAGTGCCCC 195
```

```
Qy     101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160
           |||||  ||||||||||||||||||||||||||||||||||||||  ||||||  ||  |  |||  ||
Db     196 GGGGAGACGCTCGCGCTGCGGACCGTCGGGCTGGTGCTCGGCACGCTGATCCAGTGCATC 255

Qy     161 gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg 220
           ||||||||||   ||  |||||      ||  ||                ||   |||||||||||||||||
Db     256 GACTGGGACAGAGTCGATGGCCTGGAGATTGACATGACCGCGGGTGGCGGGCTGACCATG 315

Qy     221 ccccgggccgtcccgttggaggccatgtgcagccgcgtacagctatgcgtggtgttctt 280
           |||  |||||||||||||||||||||||  |||||   |||   |||||||||||  |  |||||||
Db     316 CCCAGGGCCGTCCCGTTGGAGGCCACGTGCAAGCCTCGTGCAGCTATGCGCGATGTTCTG 375

Qy     281 aagaggctctgaaaacctcatggatcgaa 309
           |  |   |||||||    ||  |||  |   |  |
Db     376 ATGGAGCTCTGAGCCTCTGATGAAGAGTA 404


RESULT    5
BE362029
LOCUS       BE362029      654 bp     mRNA              EST        20-JUL-2000
DEFINITION  DG1_83_E01.g1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   BE362029
VERSION     BE362029.1  GI:9303586
KEYWORDS    EST.
SOURCE      sorghum.
   ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 654)
   AUTHORS  Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
            ,L.H.
   TITLE    An EST database from Sorghum: dark-grown seedlings
   JOURNAL  Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: PolyTMix
            High quality sequence start: 62
            High quality sequence stop: 654
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="Dark Grown 1 (DG1)"
                     /note="Organ: 5-day-old dark-grown seedlings; Vector:
                     Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
```

BASE COUNT   115 a   173 c   226 g   140 t
ORIGIN

    Query Match   41.4%;   Score 179;   DB 10;   Length 654;
    Best Local Similarity   78.8%;   Pred. No. 2.4e-33;
    Matches 212; Conservative   0; Mismatches   57; Indels   0; Gaps   0;

Qy   41 ggcggcaaggccaagggggcccctgctgatccctttcgggatggggcggcccaattgcccc 100
           | |||||||||||| ||||||| |||||||| || |||||||||||||||||| ||| |||||||
Db   196 GACGGCAAGGCCGAGGGGCGGCTGCTGATGCCGTTCGGGATGGGGCGGCGCAAGTGCCCC 255

Qy   101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160
           ||||| |||||||||||||||||||||||||||||||||||||| |||||| || | ||| ||
Db   256 GGGGAGACGCTCGCGCTGCGGACCGTCGGGCTGGTGCTCGGCACGCTGATCCAGTGCATC 315

Qy   161 gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg 220
           |||||||||| || |||||| || || || ||||||||||||||||
Db   316 GACTGGGACAGAGTCGATGGCCTGGAGATTGACATGACCGCGGGTGGCGGGCTGACCATG 375

Qy   221 ccccgggccgtcccgttggaggccatgtgcagccgcgtacagctatgcgtggtgttctt 280
           ||| ||||||||||||||||||||||| |||||| ||| ||| |||||||||| | |||||||
Db   376 CCCAGGGCCGTCCCGTTGGAGGCCACGTGCAAGCCTCGTGCAGCTATGCGCGATGTTCTG 435

Qy   281 aagaggctctgaaaacctcatggatcgaa 309
           | | ||||||| || ||| | | |
Db   436 ATGGAGCTCTGAGCCTCTGATGAAGAGTA 464

RESULT   6
AW676742
LOCUS   AW676742   693 bp   mRNA   EST   19-JUL-2000
DEFINITION DG1_14_A08.g1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   AW676742
VERSION   AW676742.1 GI:7550409
KEYWORDS   EST.
SOURCE   sorghum.
  ORGANISM Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1 (bases 1 to 693)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
            ,L.H.
  TITLE   An EST database from Sorghum: dark-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            ·Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions

below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: T7
High quality sequence start: 76
High quality sequence stop: 693
POLYA=Yes.
FEATURES   Location/Qualifiers
     source   1. .693
                  /organism="Sorghum bicolor"
                  /db_xref="taxon:4558"
                  /clone_lib="Dark Grown 1 (DG1)"
                  /note="Organ: 5-day-old dark-grown seedlings; Vector:
                  Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
                  made from poly-A RNA in the cloning vector lambda ZAP II.
                  Clones to be sequenced were prepared by mass excision."
BASE COUNT   124 a   180 c   231 g   158 t
ORIGIN

   Query Match   41.4%;   Score 179;   DB 10;   Length 693;
   Best Local Similarity   78.8%;   Pred. No. 2.4e-33;
   Matches 212; Conservative   0; Mismatches   57; Indels   0; Gaps   0;

Qy   41 ggcggcaaggccaagggggcccctgctgatccctttcgggatggggcggcccaattgcccc 100
           | |||||||||| |||||| ||||||||| || |||||||||||||||| ||| ||||||
Db   189 GACGGCAAGGCCGAGGGGCGGCTGCTGATGCCGTTCGGGATGGGGCGGCGCAAGTGCCCC 248

Qy   101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160
           ||||| |||||||||||||||||||||||||||||||||||||| |||||| || | ||| ||
Db   249 GGGGAGACGCTCGCGCTGCGGACCGTCGGGCTGGTGCTCGGCACGCTGATCCAGTGCATC 308

Qy   161 gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg 220
           ||||||||||| || ||||| || || || |||||||||||||||||
Db   309 GACTGGGACAGAGTCGATGGCCTGGAGATTGACATGACCGCGGGTGGCGGGCTGACCATG 368

Qy   221 ccccgggccgtcccgttggaggccatgtgcagccgcgtacagctatgcgtggtgttctt 280
           ||| ||||||||||||||||||||||||| |||||| ||| ||| ||||||||||| | |||||||
Db   369 CCCAGGGCCGTCCCGTTGGAGGCCACGTGCAAGCCTCGTGCAGCTATGCGCGATGTTCTG 428

Qy   281 aagaggctctgaaaacctcatggatcgaa 309
           | | ||||||| || ||| | | |
Db   429 ATGGAGCTCTGAGCCTCTGATGAAGAGTA 457

RESULT   7
BE357860
LOCUS   BE357860   695 bp   mRNA   EST   20-JUL-2000
DEFINITION   DG1_22_C12.g1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
             sequence.
ACCESSION   BE357860
VERSION   BE357860.1 GI:9299417
KEYWORDS   EST.
SOURCE   sorghum.
  ORGANISM   Sorghum bicolor
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
             clade; Panicoideae; Andropogoneae; Sorghum.

FEATURES   Location/Qualifiers
     source   1. .695
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="Dark Grown 1 (DG1)"
                     /note="Organ: 5-day-old dark-grown seedlings; Vector:
                     Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
                     made from poly-A RNA in the cloning vector lambda ZAP II.
                     Clones to be sequenced were prepared by mass excision."
BASE COUNT   121 a   197 c   247 g   130 t
ORIGIN

   Query Match   41.4%; Score 179; DB 10; Length 695;
   Best Local Similarity 78.8%; Pred. No. 2.4e-33;
   Matches 212; Conservative   0; Mismatches 57; Indels   0; Gaps   0;

Qy   41 ggcggcaaggccaagggggcccctgctgatccctttcgggatggggcggcccaattgcccc 100
          | ||||||||||| |||||| |||||||| || ||||||||||||||||||| ||| ||||||
Db   286 GACGGCAAGGCCGAGGGGCGGCTGCTGATGCCGTTCGGGATGGGGCGGCGCAAGTGCCCC 345

Qy   101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160
          ||||| ||||||||||||||||||||||||||||||||||||||| |||||| || | ||| ||
Db   346 GGGGAGACGCTCGCGCTGCGGACCGTCGGGCTGGTGCTCGGCACGCTGATCCAGTGCATC 405

Qy   161 gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg 220
          |||||||||| || ||||| || || || || || ||||||||||||||
Db   406 GACTGGGACAGAGTCGATGGCCTGGAGATTGACATGACCGCGGGTGGCGGGCTGACCATG 465

Qy   221 ccccgggccgtcccgttggaggccatgtgcagccgcgtacagctatgcgtggtgttctt 280
          ||| ||||||||||||||||||||||| |||||| ||| ||| ||||||||| | ||||||
Db   466 CCCAGGGCCGTCCCGTTGGAGGCCACGTGCAAGCCTCGTGCAGCTATGCGCGATGTTCTG 525

Qy   281 aagaggctctgaaaacctcatggatcgaa 309
          | | |||||||| || ||| | | |
Db   526 ATGGAGCTCTGAGCCTCTGATGAAGAGTA 554

RESULT   8
AW922289
LOCUS   AW922289   535 bp   mRNA   EST   19-JUL-2000
DEFINITION   DG1_17_H09.g1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
             sequence.
ACCESSION   AW922289
VERSION   AW922289.1 GI:8088114
KEYWORDS   EST.
SOURCE   sorghum.
  ORGANISM   Sorghum bicolor
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
             clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1 (bases 1 to 535)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
             ,L.H.
  TITLE   An EST database from Sorghum: dark-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Cordonnier-Pratt MM
             Department of Botany
             The University of Georgia
             Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
             Tel: 706 542 1860
             Fax: 706 542 1805
             Email: mmpratt@uga.edu
             Sequences have been trimmed to exclude PolyA, vector and regions
             below Phred quality 16. The threshold for highest quality sequence
             is 20.
             Seq primer: PolyTMix
             High quality sequence start: 10
             High quality sequence stop: 511
             POLYA=No.
FEATURES   Location/Qualifiers
     source   1. .535
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="Dark Grown 1 (DG1)"
                     /note="Organ: 5-day-old dark-grown seedlings; Vector:
                     Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
                     made from poly-A RNA in the cloning vector lambda ZAP II.
                     Clones to be sequenced were prepared by mass excision."
BASE COUNT   109 a   113 c   163 g   150 t
ORIGIN

   Query Match   39.2%;   Score 169.2; DB 10; Length 535;
   Best Local Similarity   78.5%;   Pred. No. 5.6e-31;
   Matches 201; Conservative   0; Mismatches   55; Indels   0; Gaps   0;

Qy   54 aggggcccctgctgatccctttcgggatggggcggcccaattgccccgggggaaacgctcg 113
           |||||| |||||||||| || ||||||||||||||||| ||| |||||||||||| |||||||
Db   1 AGGGGCGGCTGCTGATGCCGTTCGGGATGGGGCGGCGCAAGTGCCCCGGGGAGACGCTCG 60

Qy   114 cgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttcgactgggacacgg 173
           |||||||||||||||||||||||||||||| ||||||| || | ||| ||||||||||| |
Db   61 CGCTGCGGACCGTCGGGCTGGTGCTCGGCACGCTGATCCAGTGCATCGACTGGGACAGAG 120

```
Qy     174 ttgatggagctcaggtttgacatgaagctancggcgggctgaccatgccccgggccgtcc 233
               |  |||||       || ||          ||     |||||||||||||||||||| |||||||||
Db     121 TCGATGGCCTGGAGATTGACATGACCGCGGGTGGCGGGCTGACCATGCCCAGGGCCGTCC 180

Qy     234 cgttggaggccatgtgcangccgcgtacagctatgcgtggtgttcttaagaggctctgaa 293
               |||||||||||| ||||| ||| ||| |||||||||| | |||||| | |    |||||||
Db     181 CGTTGGAGGCCACGTGCAAGCCTCGTGCAGCTATGCGCGATGTTCTGATGGAGCTCTGAG 240

Qy     294 aacctcatggatcgaa 309
               || ||| |   | |
Db     241 CCTCTGATGAAGAGTA 256


RESULT    9
AI668207/c
LOCUS       AI668207       584 bp      mRNA              EST        02-FEB-2000
DEFINITION  605018C02.x1 605 - Endosperm cDNA library from Schmidt lab Zea mays
            cDNA, mRNA sequence.
ACCESSION   AI668207
VERSION     AI668207.1  GI:4827515
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 584)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 605018  row: C  column: 02.
FEATURES             Location/Qualifiers
     source          1. .584
                     /organism="Zea mays"
                     /cultivar="Ohio43"
                     /db_xref="taxon:4577"
                     /clone_lib="605 - Endosperm cDNA library from Schmidt lab"
                     /tissue_type="nucellar, embryo, and endosperm"
                     /dev_stage="10-14 days post-pollination"
                     /lab_host="DH5(alpha)"
                     /note="Organ: Kernel; Vector: pAD-GAL4-2'; Site_1: EcoRI;
                     Site_2: XhoI; Kernel endosperm cDNA library from Schmidt
                     lab"
BASE COUNT       111 a      188 c      183 g      102 t
ORIGIN


   Query Match              37.7%;  Score 163;  DB 10;  Length 584;
   Best Local Similarity    77.5%;  Pred. No. 1.8e-29;
```

```
Matches  196; Conservative   0;  Mismatches  57;  Indels   0;  Gaps    0;

Qy     40 cggcggcaaggccaaggggcccctgctgatccctttcgggatggggcggcccaattgccc 99
          || |||||||||| |||| |  ||| || | || |||||||||||| |||| ||  |||||
Db    391 CGACGGCAAGGCCGAGGGCCGGCTGATGCTGCCGTTCGGGATGGGACGGCGCAGGTGCCC 332

Qy    100 cggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgctt 159
          ||||||  ||  ||||||||||||||||| ||| || ||||||||| |||||  || ||||
Db    331 CGGGGAGACACTCGCGCTGCGGACCGCCGGCCTCGTGCTCGCCACGCTCATCCAGTGCTT 272

Qy    160 cgactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccat 219
          | |||||||| | | ||||| ||| || |                | | |||||||| |||||
Db    271 CCACTGGGACAGGATCGATGGCGCTGAGATCGACATGACCGAGAGCGGCGGGCTCACCAT 212

Qy    220 gccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatgcgtggtgttct 279
          ||||||||||||||||||||||||||||||| |||| ||| ||   ||| |||||||  ||||||
Db    211 GCCCCGGGCCGTCCCGTTGGAGGCCACCTGCAAGCCTCGCGAAGCCATGCGTCATGTTCT 152

Qy    280 taagaggctctga 292
          | ||  |||||||
Db    151 TCAGCAGCTCTGA 139


RESULT  10
BG320973/c
LOCUS       BG320973      790 bp    mRNA         EST       27-FEB-2001
DEFINITION  Zm04_02d03_A Zm04_AAFC_ECORC_cold_stressed_maize_seedlings Zea mays
            cDNA clone Zm04_02d03, mRNA sequence.
ACCESSION   BG320973
VERSION     BG320973.1  GI:13150651
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 790)
  AUTHORS   Singh,J.A., Wakui,K., Couroux,P., De Moors,A., Harris,L.J., Hattori
            ,J.I., Ouellet,T., Robert,L.S., Sprott,D. and Tinker,N.A.
    TITLE   Expressed Sequence Tags from Cold-Stressed Maize Seedlings
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Singh,J.A.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            960 Carling Avenue, Bldg. 20, Ottawa, Ontario, K1A 0C6, Canada
            Tel: (613) 759-1662
            Fax: (613) 759-1701
            Email: singhja@em.agr.ca.
FEATURES             Location/Qualifiers
     source          1. .790
                     /organism="Zea mays"
                     /cultivar="CO328"
                     /db_xref="taxon:4577"
                     /clone="Zm04_02d03"
                     /clone_lib="Zm04_AAFC_ECORC_cold_stressed_maize_seedlings"
                     /tissue_type="Leaf, crown"
                     /note="Vector: Bluescript SK-/XhoI-EcoRI; Site_1: Eco RI;
```

BASE COUNT   145 a   256 c   241 g   143 t   5 others
ORIGIN

     Query Match   34.8%;   Score 150.4; DB 11;   Length 790;
     Best Local Similarity   76.8%;   Pred. No. 1.9e-26;
     Matches 195; Conservative   0; Mismatches   58; Indels   1; Gaps   1;

```
Qy     40 cggcggcaaggccaagggggcccctgctgatccctttcgggatggggcggcccaattgccc 99
          || |||||||||| |||| |  ||| || | || |||||||||||| |||| ||  |||||
Db    393 CGACGGCAAGGCCGAGGGCCGGCTGATGCTGCCGTTCGGGATGGGACGGCGCAGGTGCCC 334

Qy    100 cggggaaacg-ctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgct 158
          |||||| ||| |||||| |||||||||| ||| || ||||||||| |||||| || | ||||
Db    333 CGGGGAGACGCCTCGCTCTGCGGACCGCCGGCCTCGTGCTCGCCACGCTCATCCAGTGCT 274

Qy    159 tcgactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgacca 218
          || |||||||||   |  |||||| ||| || |  |        |  | |||||||||| ||||
Db    273 TCCACTGGGACAGAATCGATGGCGCTGAGATCGACATGACCGAGAGCGGCGGGCTCACCA 214

Qy    219 tgccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatgcgtggtgttc 278
          |||||||||||||||||||||||||||||| |||| ||| ||    ||| |||||| |||||
Db    213 TGCCCCGGGCCGTCCCGTTGGAGGCCACCTGCAAGCCTCGCGAAGCCATGCGTCATGTTC 154

Qy    279 ttaagaggctctga 292
          || || ||||||||
Db    153 TTCAGCAGCTCTGA 140
```


RESULT   11
BG464759
LOCUS   BG464759   357 bp   mRNA   EST   20-MAR-2001
DEFINITION EM1_33_G05.g1_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   BG464759
VERSION   BG464759.1 GI:13393586
KEYWORDS   EST.
SOURCE   sorghum.
  ORGANISM Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1 (bases 1 to 357)
  AUTHORS   Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H.
  TITLE   An EST database from Sorghum: developing embryos
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805

Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: PolyTMix
High quality sequence start: 101
High quality sequence stop: 228
POLYA=No.

FEATURES   Location/Qualifiers
    source   1. .357
                      /organism="Sorghum bicolor"
                      /db_xref="taxon:4558"
                      /clone_lib="Embryo 1 (EM1)"
                      /note="Organ: Embryos germinated for 24 hr; Vector:
                      pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2:
                      EcoRI; The library was made from poly-A RNA in the cloning
                      vector lambda ZAP II. Clones to be sequenced were
                      prepared by mass excision."
BASE COUNT   63 a   86 c   127 g   79 t   2 others
ORIGIN

  Query Match   34.8%;   Score 150.2;   DB 11;   Length 357;
  Best Local Similarity   66.7%;   Pred. No. 2.2e-26;
  Matches 226; Conservative   0; Mismatches 112; Indels   1; Gaps   1;

Qy   78 ggatggggcggcccaattgccccggggaaacgctcgcgctgcggaccgtcgggctggtgc 137
           ||| ||||||||| ||| ||||||||||| ||||||||||||||||||||||||||||||
Db   1 GGAAGGGGCGGCGCAAGTGCCCCGGGGAGACGCTCGCGCTGCGGACCGTCGGGCTGGTGC 60

Qy   138 tcgcaacgctgctcaattgcttcgactgggacacggttgatggagctcaggtttgacatg 197
           ||| |||||| || | ||| ||||||||||| || || || || ||
Db   61 TCGGCACGCTGATCCAGTGCATCGACTGGGACAGAGTCGANGGCCTGGAGATTGACATGA 120

Qy   198 aagctancggcgggctgaccatgccccgggccgtcccgttggaggccatgtgcangccgc 257
             || ||||||||||||||||||| ||||||||||||||||||||| ||||| ||| |
Db   121 CCGCGGGTGGCGGGCTGACCATGCCCAGGGCCGTCCCGTTGGAGGCCACGTGCAAGCCTC 180

Qy   258 gtacagctatgcgtggtgttcttaagaggctctgaaaacctcatggatcgaattgctggc 317
           || ||||||||||| | |||||| | | ||||||| || | | ||||
Db   181 GTGCAGCTATGCGCGATGTTCTGATGGAGCTCTGAGCCTCTATGAAGAGTACATCTTGGC 240

Qy   318 -atcgtctgaagggtgtatgacgtagcttccgagttccgagcatatatattcacttgcct 376
            || || ||||| | | | ||| | | | || | | |
Db   241 AATGATCCCTAGGGTCTCACTGCGTGGTACTGAGGTTCAACCGGTACTAGTGTGTAGGTG 300

Qy   377 tgtaacaatttaattttcgccgattgtatggaatggatt 415
           |||| || | ||| || || || || || | |
Db   301 TGTAGCAGTANTGCTTTGGCTTATGGTGTGTGCTGAACT 339

RESULT 12
BG464902
LOCUS   BG464902   293 bp   mRNA   EST   20-MAR-2001
DEFINITION   EM1_35_G06.g1_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA
             sequence.
ACCESSION   BG464902

```
VERSION      BG464902.1  GI:13393837
KEYWORDS     EST.
SOURCE       sorghum.
  ORGANISM   Sorghum bicolor
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
             clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE    1  (bases 1 to 293)
  AUTHORS    Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H.
  TITLE      An EST database from Sorghum: developing embryos
  JOURNAL    Unpublished (2000)
COMMENT      Contact: Cordonnier-Pratt MM
             Department of Botany
             The University of Georgia
             Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
             Tel: 706 542 1860
             Fax: 706 542 1805
             Email: mmpratt@uga.edu
             Sequences have been trimmed to exclude PolyA, vector and regions
             below Phred quality 16. The threshold for highest quality sequence
             is 20.
             Seq primer: PolyTMix
             High quality sequence start: 7
             High quality sequence stop: 219
             POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .293
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="Embryo 1 (EM1)"
                     /note="Organ: Embryos germinated for 24 hr; Vector:
                     pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2:
                     EcoRI; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
BASE COUNT       51 a      79 c     104 g      58 t        1 others
ORIGIN
```

```
  Query Match            34.4%;   Score 148.6;  DB 11;   Length 293;
  Best Local Similarity  77.4%;   Pred. No. 5.5e-26;
  Matches 178;  Conservative   0;  Mismatches   52;  Indels   0;  Gaps    0;

Qy     80 atggggcggcccaattgccccgggggaaacgctcgcgctgcggaccgtcgggctggtgctc 139
          ||||||||||| ||| ||||||||||| |||||||||||||||||||||||||||||||||
Db      1 ATGGGGCGGCGCAAGTGCCCCGGGGAGACGCTCGCGCTGCGGACCGTCGGGCTGGTGCTC 60

Qy    140 gcaacgctgctcaattgcttcgactgggacacggttgatggagctcaggtttgacatgaa 199
          |   ||||||| || | ||| ||||||||||||| || |||||     || ||
Db     61 GGCACGCTGATCCAGTGCATCGACTGGGACAGAGTCGATGGCCTGGAGATTGACATGACC 120

Qy    200 gctancggcgggctgaccatgccccgggccgtcccgttggaggccatgtgcangccgcgt 259
          ||       ||||||||||||||||||| |||||||||||||||||||||| ||||| ||| |||
Db    121 GCGGGTGGCGGGCTGACCATGCCCAGGGCCGTCCCGTTGGAGGCCACGTGCAAGCCTCGT 180

Qy    260 acagctatgcgtggtgttcttaagaggctctgaaaacctcatggatcgaa 309
          |||||||||| | |||||| | |  ||||||   || | | | | ||
```

Db   181 GCAGCTATGCGCGATGTTCTGATGGAGCTCTGAGCCTCTGANGAAGAGTA 230

RESULT 13
BE704790/c
LOCUS   BE704790   803 bp   mRNA   EST   12-SEP-2000
DEFINITION   Sc02_02e10_A Sc02_AAFC_ECORC_cold_stressed_winter_rye_seedlings
                Secale cereale cDNA clone Sc02_02e10, mRNA sequence.
ACCESSION   BE704790
VERSION   BE704790.1 GI:10093055
KEYWORDS   EST.
SOURCE   rye.
  ORGANISM   Secale cereale
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
                ; Triticeae; Secale.
REFERENCE   1 (bases 1 to 803)
  AUTHORS   Singh,J.A., Piche,C., Couroux,P., De Moors,A., Harris,L.J., Hattori
                ,J.I., Ouellet,T., Robert,L.S., Sprott,D. and Tinker,N.A.
  TITLE   Expressed Sequence Tags from Cold-Stressed Winter Rye Seedlings
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Singh,J.A.
                Eastern Cereal and Oilseed Research Centre
                Agriculture and Agri-food Canada
                960 Carling Avenue, Bldg. 20, Ottawa, Ontario, K1A 0C6, Canada
                Tel: (613) 759-1662
                Fax: (613) 759-1701
                Email: singhja@em.agr.ca.
FEATURES   Location/Qualifiers
     source   1. .803
                     /organism="Secale cereale"
                     /cultivar="Puma (winter rye)"
                     /db_xref="taxon:4550"
                     /clone="Sc02_02e10"
                     /clone_lib="Sc02_AAFC_ECORC_cold_stressed_winter_rye_seedl
                     ings"
                     /tissue_type="leaf, crown"
                     /dev_stage="seedling three-leaf stage"
                     /note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: Eco RI;
                     Site_2: Xho I; Sampled three-leaf seedlings treated for
                     one week at 2oC, 12 hrs light/day. Library made with
                     Stratagene UNIZAP XR Kit/(not packaged). cDNA is directly
                     ligated into SK+/XhoI-EcoRI, then electroporated into
                     TOP10 cells (Invitrogen)."
BASE COUNT   157 a   279 c   221 g   143 t   3 others
ORIGIN

   Query Match   33.8%;   Score 145.8; DB 10;   Length 803;
   Best Local Similarity   72.4%;   Pred. No. 2.4e-25;
   Matches 186; Conservative   1; Mismatches   70; Indels   0; Gaps   0;

Qy   41 ggcggcaaggccaaggggcccctgctgatccctttcgggatggggcggcccaattgcccc 100
           | ||||||||||| ||||||| | ||||||| |||||||||||| |||| ||||||
Db   494 GACGGCAAGGCCGAGGGGCGGTTCATGATCCCGTTCGGGATGGGCCGGCGGCGGTGCCCC 435

Qy   101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160

```
           |||||  |:|||  |||||||||||| ||||  ||||||| || ||||||| |  | |||||
Db    434  GGGGAGAMGCTGGCGCTGCGGACCATCGGCATGGTGCTGGCCACGCTGGTGCAGTGCTTT  375

Qy    161  gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg  220
           |||||||||    ||  |||||| ||  ||||          |        |||||||||| |||||
Db    374  GACTGGGAGCGCGTGGATGGCGCGGAGGTGGACATGACGGAGGGCGGCGGGCTCACCATC  315

Qy    221  ccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatgcgtggtgttctt  280
           |||  |||| | ||| | ||||||| |||||| |||||| || || |||||| |  || |||
Db    314  CCCAAGGCCATGCCGCTTGAGGCCGTGTGCAGGCCGCGCACGGCCATGCGCGACGTGCTT  255

Qy    281  aagaggctctgaaaacc  297
           |||| ||||||    ||
Db    254  CAGAGCCTCTGATGGCC  238


RESULT  14
AL503532/c
LOCUS       AL503532       700 bp     mRNA            EST         04-JAN-2001
DEFINITION  AL503532 Hordeum vulgare Barke roots Hordeum vulgare cDNA clone
            HW02H20T 5', mRNA sequence.
ACCESSION   AL503532
VERSION     AL503532.1  GI:12029747
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   Michalek,W., Weschke,W., Pleissner,K.-P. and Graner,A.
  TITLE     EST sequencing and analysis in barley
  JOURNAL   Unpublished (2000)
  COMMENT   Contact: Michalek W
            Institute for Plant Genetics and Crop Plant Research
            Corrensstr.3, D-06466 Gatersleben, Germany
            Email: michalek@ipk-gatersleben.de, http://pgrc.ipk-gatersleben.de
            Seq primer: T3 primer for 5'end.
FEATURES             Location/Qualifiers
     source          1. .700
                     /organism="Hordeum vulgare"
                     /cultivar="Barke"
                     /db_xref="taxon:4513"
                     /clone="HW02H20T"
                     /clone_lib="Hordeum vulgare Barke roots"
                     /tissue_type="roots"
                     /lab_host="XLOLR"
                     /note="Vector: plasmid pBK-CMV; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was made from roots of spring barley variety
                     'Barke', a high quality malting variety. Roots were grown
                     for two days on filter paper at room temperature Cloning
                     sites: EcoRI (5'-end of cDNA) and XhoI (3'-end of cDNA).
                     NOTE: Due to a cloning artefact caused by the kit, in most
                     cases the EcoRI site is NOT present, as well as the EcoRI
                     adapter. Average insert size is 1 kb Sequence trimming:
                     Vector sequences and sequence ends were trimmed from the
                     5'-and 3'-end until a 50 bp window contains less than two
```

BASE COUNT   132 a   229 c   196 g   137 t   6 others
ORIGIN

   Query Match   33.0%;   Score 142.4;   DB 10;   Length 700;
   Best Local Similarity   72.2%;   Pred. No. 1.6e-24;
   Matches 182;   Conservative   0;   Mismatches   70;   Indels   0;   Gaps   0;

```
Qy    41 ggcggcaaggccaagggggcccctgctgatccctttcgggatggggcggcccaattgcccc 100
         I IIIIIIIIII IIIIII    I   IIIIIIII IIIIIIIIIII II I      IIIIII
Db   476 GACGGCAAGGCCGAGGGGCGGTTCATGATCCCGTTCGGGATGGGCCGCCGGCGGTGCCCC 417

Qy   101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160
         IIIII IIIII IIIIIIIIIIII IIII  IIIIIII II IIIIII I  I IIIIII
Db   416 GGGGANACGCTGGCGCTGCGGACCATCGGCATGGTGCTGGCCACGCTGGTGCAGTGCTTC 357

Qy   161 gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg 220
         IIIIIIIII   II II II    IIII       I  I IIIIIIIIII IIIII
Db   356 GACTGGGACCGCGTCGACGGCAAGGAGGTGGACATGACGGAGAGCGGCGGGCTCACCATC 297

Qy   221 ccccgggccgtcccgttggaggccatgtgcagccgcgtacagctatgcgtggtgttctt 280
         III  IIIIIII III I IIIIIII I IIII IIIIII I II IIIIII I  II II
Db   296 CCCAAGGCCGTGCCGCTCGAGGCCGTNTGCAGGCCGCGCCCGGCCATGCGCGACGTGCTC 237

Qy   281 aagaggctctga 292
         IIII IIIIII
Db   236 CAGAGCCTCTGA 225
```


RESULT   15
BE412662/c
LOCUS   BE412662   702 bp   mRNA   EST   24-JUL-2000
DEFINITION   MCG007.D10R990625 ITEC MCG Barley Leaf/Culm Library Hordeum vulgare
                cDNA clone MCG007.D10, mRNA sequence.
ACCESSION   BE412662
VERSION   BE412662.1 GI:9410620
KEYWORDS   EST.
SOURCE   barley.
  ORGANISM   Hordeum vulgare
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
                ; Triticeae; Hordeum.
REFERENCE   1 (bases 1 to 702)
  AUTHORS   Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
                ,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
                Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
                Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
                Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
                Sorrells,M., Warburton,M. and Wenzel,G.
  TITLE   International Triticeae EST Cooperative (ITEC): Production of
                Expressed Sequence Tags for Species of the Triticeae
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Graner A
                Institute for Plant Genetics & Crop Plant Research
                Corrensstr. 3, D-06466 Gatersleben GERMANY
                Tel: 49 39482 5521

FEATURES   Location/Qualifiers
    source   1. .702
                   /organism="Hordeum vulgare"
                   /db_xref="taxon:4513"
                   /clone="MCG007.D10"
                   /clone_lib="ITEC MCG Barley Leaf/Culm Library"
                   /tissue_type="leaf/culm"
                   /dev_stage="etiolated"
BASE COUNT   132 a   229 c   198 g   137 t   6 others
ORIGIN

   Query Match   33.0%;   Score 142.4; DB 10;   Length 702;
   Best Local Similarity 72.2%;   Pred. No. 1.6e-24;
   Matches 182; Conservative   0; Mismatches   70; Indels   0; Gaps   0;

Qy   41 ggcggcaaggccaaggggcccctgctgatccctttcgggatggggcggcccaattgcccc 100
          | |||||||||| |||||| | ||||||| |||||||||| || | ||||||
Db   476 GACGGCAAGGCCGAGGGGCGGTTCATGATCCCGTTCGGGATGGGCCGCCGGCGGTGCCCC 417

Qy   101 ggggaaacgctcgcgctgcggaccgtcgggctggtgctcgcaacgctgctcaattgcttc 160
          ||||| ||||| |||||||||||||| |||| ||||||| || ||||||| | | ||||||
Db   416 GGGGANACGCTGGCGCTGCGGACCATCGGCATGGTGCTGGCCACGCTGGTGCAGTGCTTC 357

Qy   161 gactgggacacggttgatggagctcaggtttgacatgaagctancggcgggctgaccatg 220
          ||||||||| || || || |||| | | ||||||||||| |||||
Db   356 GACTGGGACCGCGTCGACGGCAAGGAGGTGGACATGACGGAGAGCGGCGGGCTCACCATC 297

Qy   221 ccccgggccgtcccgttggaggccatgtgcangccgcgtacagctatgcgtggtgttctt 280
          ||| ||||||| |||| | ||||||| | |||| ||||||| | || ||||| | | ||
Db   296 CCCAAGGCCGTGCCGCTCGAGGCCGTNTGCAGGCCGCGCCCGGCCATGCGCGACGTGCTC 237

Qy   281 aagaggctctga 292
           |||| |||||||
Db   236 CAGAGCCTCTGA 225

Search completed: February 7, 2002, 08:20:54
Job time: 18131 sec